Earthquakes of the Nepal Himalaya: towards a physical model of the seismic cycle

Thesis by
Thomas J. Ader

In Partial Fulfillment of the Requirements for the degree
of
Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California
2013
(Defended May 21st 2013)

If . . .

Rudyard Kipling

*“If you can keep your head when all about you
Are losing theirs and blaming it on you;
If you can trust yourself when all men doubt you,
But make allowance for their doubting too:
If you can wait and not be tired by waiting,
Or being lied about, don’t deal in lies,
Or being hated don’t give way to hating,
And yet don’t look too good, nor talk too wise;*

*If you can dream — and not make dreams your master;
If you can think — and not make thoughts your aim,
If you can meet with Triumph and Disaster
And treat those two impostors just the same:
If you can bear to hear the truth you’ve spoken
Twisted by knaves to make a trap for fools,
Or watch the things you gave your life to, broken,
And stoop and build ’em up with worn-out tools;*

*If you can make one heap of all your winnings
And risk it on one turn of pitch-and-toss,
And lose, and start again at your beginnings
And never breathe a word about your loss:
If you can force your heart and nerve and sinew
To serve your turn long after they are gone,
And so hold on when there is nothing in you
Except the Will which says to them: ‘Hold on!’*

*If you can talk with crowds and keep your virtue,
Or walk with Kings — nor lose the common touch,
If neither foes nor loving friends can hurt you,
If all men count with you, but none too much:
If you can fill the unforgiving minute
With sixty seconds worth of distance run,
Yours is the Earth and everything that’s in it,
And — which is more — you’ll be a Man, my son!”*

Acknowledgements

A PhD is a significant chunk of life. The journey has been more arduous than I had initially assumed, but also much richer and, above everything else, unexpected. If I have to make one wish for the future, it is that it may remain as unexpected and fulfilling as the five years I have had the incredible joy to spend here. I came here with fervor, high expectations, and their counterpart of doubts. One doubt concerned leaving aside my childhood passion for astronomy to study the only planet that does not revolve above our heads: the Earth. Luckily, my advisor, Jean-Philippe, found the perfect compromise and sent me mountain-gazing in the Himalaya, where more than any other place on the globe, admiring the scenery requires to look up. I would therefore like to start this section by extending my deepest gratitude to him, for giving me the opportunity to work on such an incredible project and make my doctoral work a true adventure, which breadth and teachings went far beyond the scientific sphere. During my entire PhD, Jean-Philippe has always displayed an unfailing and highly contagious enthusiasm, providing me with the guidance, encouragement and support I needed. His permanent availability, together with the undivided attention and positive outlook that he has always granted to my work, have been instrumental in the smooth progress of my research. He trusted me to organize my work schedule in a way that would make my time here as enjoyable as possible, and let me orient my research in the direction that looked the most appealing to me.

This work would not have been possible without the multiple collaborations I have felt privileged to establish here at Caltech. Each of the four chapters of this thesis is a paper, either published, submitted or in the last stages of preparation, and the co-authors were an integral part of the work. Jean-Paul Ampuero selflessly shared with me the fertile idea that blossomed into chapter 3, and thoroughly assisted me in bringing it to maturation in just a few months. I probably owe millions of dollars in computer money to Nadia Lapusta, who let me use the BICYCLE algorithm she wrote for the simulations of chapter 4. I feel greatly *indebted toward her for that and for all the time she took from her busy schedule to share with me her insights, encouragement, and help for me to develop chapter 4. Thanks also to Mark Simons, the fourth musketeer of my thesis committee, for his helpful comments on the manuscript.*

Although more than 6000 miles away, I have been lucky enough to be given the opportunity to maintain tight bonds with France during my PhD. First of all, I would like to acknowledge the precious financial help from the Ministère de la Recherche and the École Polytechnique for the three-year funding they granted me. Hélène Lyon-Caen was my advisor at the ENS, and played a big role in the achievement of this thesis through her helpful comments and suggestions, and by having made the collaboration between Caltech and the ENS possible. Laurent Bollinger unsparingly took time from his schedule as soon as I needed advice on Nepalese seismicity catalogs or an excuse to fly back home, answering my questions in depth and with an utter passion, sharing with me all his expertise from the current seismicity of Nepal to ancient hidden discoveries. Our lengthy and fruitful discussions have been a huge source of inspiration in the development of my research and have coincidently drawn my attention on the political hurdles that can stand in the way of a scientific project like the one I was working on.

From the anarchic cacophony of Kathmandu to the peaceful wisdom of the remote Mustang kingdom, each of my field trips in Nepal has been a unique and breath-taking experience forever engraved in my memory. These would not have been possible without the supernatural planning skills of John Galetzka, the inspiring kind of person that nothing seems able to stop, who always knows what to do, and even more spectacular, how to do it. I would also like to pay tribute to the invaluable help of the Nepalese colleagues from the National Seismological Centre in the Department of Mines and Geology in Kathmandu, who were always eager to proffer their precious assistance. Prithvi and Subhadra deserve a special mention for their kind hospitality and their always enjoyable company during my stays in Kathmandu.

Throughout this long journey far from home, my family has been an unbreakable gem of moral support for me. They are a model of cohesion for me and the main reason why I could not see myself live so far from Europe my entire life. My mother, without whom I would probably not have lasted a year here, was in particular always able to find the right words at the right time, soothing away my concerns and restoring my momentum whenever *it was necessary.*

The people I have met here are what made my five years in California so memorable, and the achievement of such a great amount of work would probably have been at the cost of my sanity if it were not for them. Out of fear of forgetting names, I will not risk mentioning each and every one of the people who made a difference in my life here. You guys are undoubtedly what I am going to miss the most when moving away. I would however like to make a special mention to Alan, who more than a dear friend has been a true mentor for me from my very first days here. Thanks Alan for teaching me how to put things into perspective, set the right priorities on my way and remain faithful to them.

Last, but not least, I would like to extend my eternal gratitude to Rudyard Kipling for gathering within a single page all the wisdom of the world, what has become the guidelines of my philosophy of life, powerful words from which I have relentlessly drawn the strength and motivation to keep my head up.

Abstract

Home to hundreds of millions of souls and land of excessiveness, the Himalaya is also the locus of a unique seismicity whose scope and peculiarities still remain to this day somewhat mysterious. Having claimed the lives of kings, or turned ancient timeworn cities into heaps of rubbles and ruins, earthquakes eerily inhabit Nepalese folk tales with the fatalistic message that nothing lasts forever. From a scientific point of view as much as from a human perspective, solving the mysteries of Himalayan seismicity thus represents a challenge of prime importance. Documenting geodetic strain across the Nepal Himalaya with various GPS and leveling data, we show that unlike other subduction zones that exhibit a heterogeneous and patchy coupling pattern along strike, the last hundred kilometers of the Main Himalayan Thrust fault, or MHT, appear to be uniformly locked, devoid of any of the creeping barriers that traditionally ward off the propagation of large events. The approximately 20 mm/yr of reckoned convergence across the Himalaya matching previously established estimates of the secular deformation at the front of the arc, the slip accumulated at depth has to somehow elastically propagate all the way to the surface at some point. And yet, neither large events from the past nor currently recorded microseismicity nearly compensate for the massive moment deficit that quietly builds up under the giant mountains. Along with this large unbalanced moment deficit, the uncommonly homogeneous coupling pattern on the MHT raises the question of whether or not the locked portion of the MHT can rupture all at once in a giant earthquake. Univocally answering this question appears contingent on the still elusive estimate of the magnitude of the largest possible earthquake in the Himalaya, and requires tight constraints on local fault properties. What makes the Himalaya enigmatic also makes it the potential source of an incredible wealth of information, and we exploit some of the oddities of Himalayan seismicity in an effort to improve the understanding of earthquake physics and cipher out the properties of the MHT. Thanks to the Himalaya, the Indo-Gangetic plain is deluged each year under a tremendous amount of water during the annual summer monsoon that collects and bears down on the Indian plate enough to pull it away from the Eurasian plate slightly, temporarily relieving a small portion of the stress mounting on the MHT. As the rainwater evaporates in the dry winter season, the plate rebounds and tension is increased back on the fault. Interestingly, the mild waggle of stress induced by the monsoon rains is about the same size as that from solid-Earth tides which gently tug at the planets solid layers, but whereas changes in earthquake frequency correspond with the annually occurring monsoon, there is no such correlation with Earth tides, which oscillate back-and-forth twice a day. We therefore investigate the general response of the creeping and seismogenic parts of MHT to periodic stresses in order to link these observations to physical parameters. First, the response of the creeping part of the MHT is analyzed with a simple spring-and-slider system bearing rate-strengthening rheology, and we show that at the transition with the locked zone, where the friction becomes near velocity neutral, the response of the slip rate may be amplified at some periods, which values are analytically related to the physical parameters of the problem. Such predictions therefore hold the potential of constraining fault properties on the MHT, but still await observational counterparts to be applied, as nothing indicates that the variations of seismicity rate on the locked part of the MHT are the direct expressions of variations of the slip rate on its creeping part, and no variations of the slip rate have been singled out from the GPS measurements to this day. When shifting to the locked seismogenic part of the MHT, spring-and-slider models with rate-weakening rheology are insufficient to explain the contrasted responses of the seismicity to the periodic loads that tides and monsoon both place on the MHT. Instead, we resort to numerical simulations using the Boundary Integral CYCLes of Earthquakes algorithm and examine the response of a 2D finite fault embedded with a rate-weakening patch to harmonic stress perturbations of various periods. We show that such simulations are able to reproduce results consistent with a gradual amplification of sensitivity as the perturbing period get larger, up to a critical period corresponding to the characteristic time of evolution of the seismicity in response to a step-like perturbation of stress. This increase of sensitivity was not reproduced by simple 1D-spring-slider systems, probably because of the complexity of the nucleation process, reproduced only by 2D-fault models. When the nucleation zone is close to its critical unstable size, its growth becomes highly sensitive to any external perturbations and the timings of produced events may therefore find themselves highly affected. A fully analytical framework has yet to be developed and further work is needed to fully describe the behavior of the fault in terms of physical parameters, which will likely provide the keys to deduce constitutive properties of the MHT from seismological observations.

Contents

Earthquakes of the Nepal Himalaya: towards a physical model of the seismic cycle

Thomas Ader, May 21st, 2013
California Institute of Technology, Pasadena, California, USA

Introduction

Whenever I tell someone that I study geophysics, I get either one of the two reactions "What is that?" or "When is the next earthquake coming?" To the first question I just reply that I study earthquakes, and face the other one within seconds. People are predictable. Earthquakes, not yet.

Having lived in Southern California for close to five years in the vicinity of a major active seismic fault, I have learned to never risk myself to a prediction more accurate than "most likely some time within the next 500 years", as many people here still live with the trauma of an intense seismic quake shaking their house, and earthquakes are a dangerous material for jokes. Adepts of conspiracy theories think that we can predict them but hide our findings for some obscure reasons, fatalist minds think that trying to predict them is as promising as trying to predict the future, and know-it-alls often explain to me that we should look for a precursor sign that would indicate that a disastrous earthquake is near.

The truth is that we are at a stage where new seismic events still raise more questions than they confirm our understanding of earthquake mechanics. The 2011 $M_W 9$ Tohoku-Oki earthquake in Japan came as a big surprise in terms of the unexpectedly huge amount of slip that occurred on the fault, generating an earthquake and a tsunami much larger than what geophysicists expected.

Five $M_W \sim 6$ earthquakes happened at Parkfield, California, with an incredibly regular 22-year period from 1881 to 1966, leaving scientists expecting an event in the late 80's. What looked like a reliable earthquake prediction turned out to be a new challenge by nature, as the following $M_W \sim 6$ event lingered for almost 20 more years past the expected date to only happen in 2004.

Earthquake prediction still slips out the grasp of current physical and statistical models and several paths are explored, hoping to find the one that will lead to the Holy Grail: being able to announce minutes, hours or even days in advance that the Earth is going to shake at a given location. We can predict sunny skies, rain, storms, snow and hurricanes days in advance using elaborate physics-based models mostly relying on fluid dynamics and meteorological observations. The key to earthquake forecasts may very well reside in similar physical models of earthquake dynamics. But the same reason that makes earthquakes so dramatic and ravaging sets them apart from the weather, this reason is their suddenness. Earthquakes happen, weather is and evolves on a human scale. This distinctive suddenness needs to be reflected in the accuracy of any physical model that could one day pretend to crack open their secret. Building such a model is no easy task and in order to do so, the entire variety of seismic configurations that the Earth has to offer needs to be exploited.

The Himalaya is one of these outstanding configurations: it is the only place on the planet where two continents collide over such a large scale. Even if the exact terminology for such a seismotectonic setting is “collision”, as it involves two continents, Figure 1 shows that it is structurally identical to a subduction such as Japan, Sumatra or Chile, which produced the largest events ever recorded during the instrumental era. Besides its paramount scientific interest, the Himalayan region is also home to hundreds of millions of people living under the ongoing threat of a devastating earthquake. The Himalaya is not a land of half measure: we know that events of magnitude greater than 8 have happened in the past and are bound to happen again. How large can Himalayan earthquake be, how often and where they should be expected remain widely debated and outstanding issues to which this doctoral work hopes to bring a valuable contribution.

More specifically, I focused my PhD on the Nepalese part of the Himalaya, where the favorable political context has permitted a collaboration with the National Seismological Center (NSC), part of the Department of Mines and Geology (DMG) in Kathmandu, Nepal. Initiated by the Département Analyse, Surveillance et Environnement (DASE) at the Commissariat à l’Energie Atomique (CEA) in France, and then extended to Caltech in California, this collaboration has seen its principal materialization in the development of both a seismic and a GPS network covering the Nepalese territory, making the country an unrivaled place to study Himalayan seismicity.

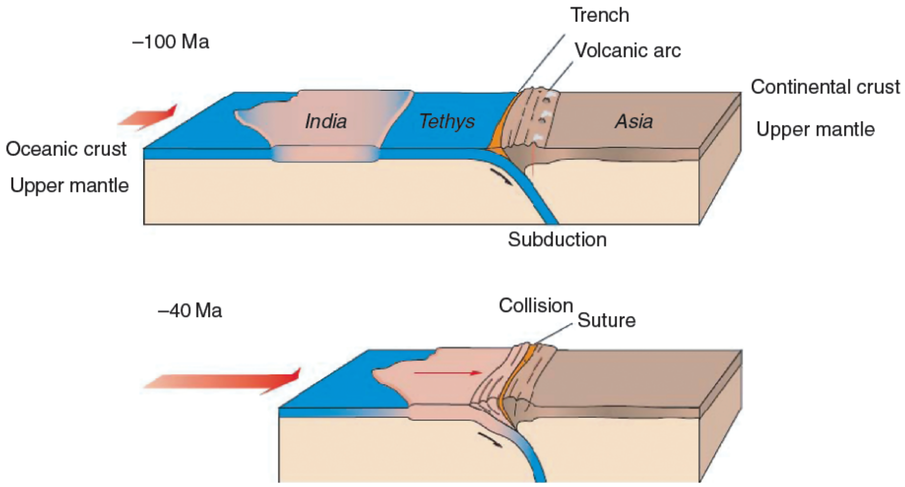


Figure 1: Formation of the Himalaya and present day collision. Prior to the collision, an ocean (the Tethys Sea) used to separate the northern margin of India and Eurasia. The southern margin of Asia was an active margin with a subduction zone similar, for example, to the Andean subduction zone bordering the western margin of South America. Today, the Indian plate is being subducted underneath the Eurasian plate and major earthquakes happen on the interface between the two plates: the Main Himalayan Thrust fault (MHT). Figure from *Avouac* (2007)

Figure 2, taken from *Ader et al.* (2012a), shows the most recent large events known to have happened within the Nepalese borders. The western part of Nepal has not ruptured since the major event of June 6th, 1505, which magnitude is still debated but most likely greater than 8 (*Ambraseys and Douglas*, 2004). The eastern part of Nepal has a more recent seismic history with an event in 1833 of magnitude slightly below 8, and the major $M_W > 8$ (e.g., *Ambraseys and Douglas*, 2004) Bihar-Nepal earthquake of 1934. Previous known events are not represented in Figure 2, but a large earthquake in 1255 is famous for having killed a third of the population of Kathmandu, among which the King of Nepal Avaya Malla, while trenches dug in western Nepal brought to light a major paleo-earthquake in $\sim$ 1100 (*Lavé et al.*, 2005), for which human reports yet await to be discovered.

1255, 1505, 1934, maybe 1100, the Himalayan collision is raising the Roof of the World at a much slower pace than the life time of collective memory, relegating the occurrence of cataclysmic earthquakes to the rank of ancient myths rather than ongoing threat. Present day survivors of the 1934 earthquake see their number vanish, taking away with them the seismic awareness in Nepal. And yet, pictures of the disaster, such as the destruction of Bhaktapur Durbar Square in Figure 3, send an alarming message to the megalopolis that Kathmandu has become today. Figure 4 is a picture that I took from the Swayambhu

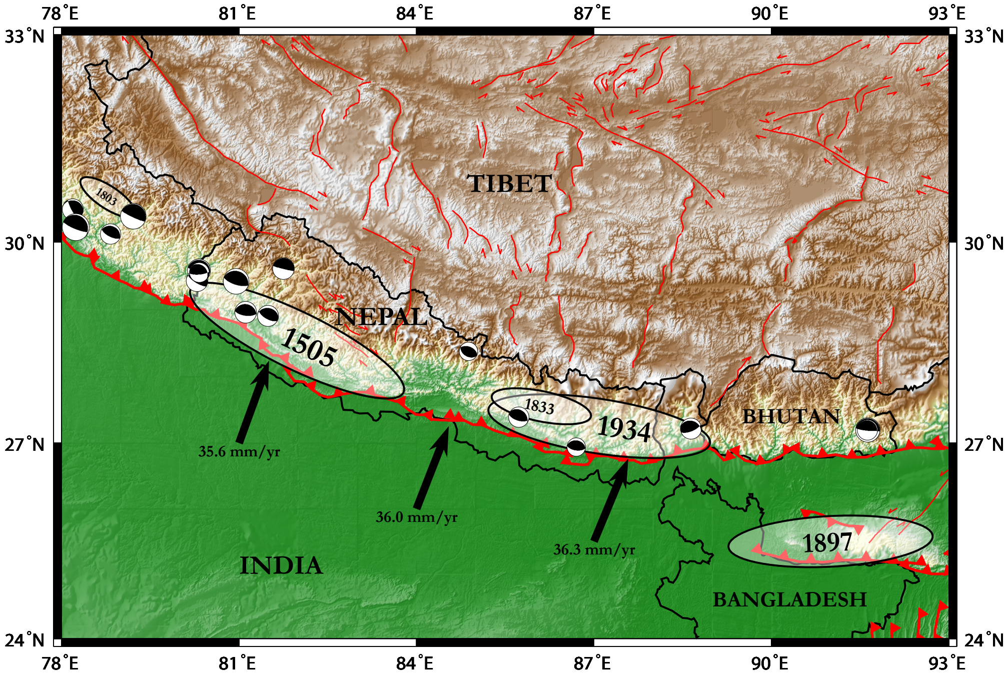


Figure 2: Seismotectonic setting of the Nepal Himalaya from *Ader et al.* (2012a). Arrows show Indian plate motion relative to Eurasia computed using the rotation poles of Eurasian plate in ITRF 2005 from *Altamimi* (2009), and Indian plate in ITRF 2005 from *Ader et al.* (2012a). Focal mechanisms show thrust events (rake = 90° ± 45°) from the CMT catalog between 1976 and 2011. White ellipses show locations of historical earthquakes according to *Ambraseys and Douglas* (2004). Ellipses sizes are scaled with the earthquakes magnitudes, and might not represent reliably the area ruptured during these earthquakes. Active faults (in red) map modified from *Styron et al.* (2011).

temple, north-east of Kathmandu, showing a city that has grown maybe too fast, spreading uncontrollably into a maze of tall and narrow 4 to 5 story buildings, essentially made of heavy bricks, an architecture well known to behave like a house of cards when undergoing a seismic tremor. The consequences of a large earthquake shaking Kathmandu today could be disastrous.

A paramount challenge in Nepal, as much from a scientific standpoint as from a human perspective, is thus to estimate how large, how often and where should major events be expected. Such information is key for establishing appropriately scaled building regulations and provides a reliable ground material to plan efficient earthquake response programs, such as the Nepal Risk Reduction Consortium Flagship projects led by the Ministry of

Figure 3: Bhaktapur Darbar Square before and after the 1934 Bihar-Nepal $M_W \sim 8.1$ earthquake that killed almost 20,000 people. Picture from *Proksch and Baidya* (1995).

Figure 4: View of Kathmandu today. The city counts close to a million inhabitants according to the latest national population census, by the National Planning Commission Secretariat, Central Bureau of Statistics (CBS), Government of Nepal in September 2011. Buildings are tall and narrow, mostly made of bricks, lining narrow streets, against all earthquake safety recommendations.

Home Affairs and the United Nations Office for the Coordination of Humanitarian Affairs (UNOCHA) [`http://un.org.np/coordinationmechanism/nrrc`].

Other major subduction zones, such as the Andean or the Sumatran subduction zones, where complete seismic cycles have been observed (i.e., slow interseismic loading / coseismic rupture / post seismic deformation), suggest that this information can be inferred from the observation of the slow interseismic loading process. Figure 5 shows that during the

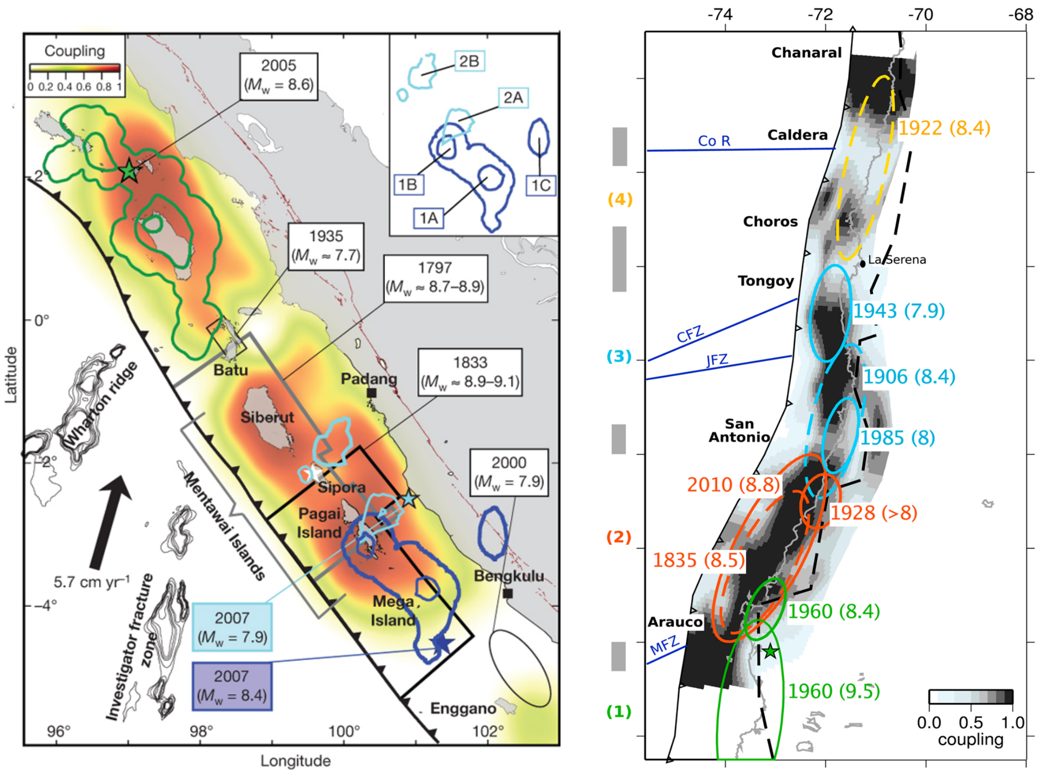


Figure 5: Coupling maps along the Sumatra (left) and Chilean (right) subduction zones, together with location of large earthquakes known to have occurred on these subduction interfaces. The color coding for the degree of coupling is different for both maps and is indicated in each map by a color bar. A coupling of 1 indicates that the fault is completely locked during the interseismic period, while a coupling of 0 corresponds to a creeping part of the fault. Known earthquakes are indicated by the contour of their slip, their year of occurrence and moment magnitude. Both the coupling map for the Sumatra (*Konca et al.*, 2008) and the Chilean subduction zones (*Métois et al.*, 2012) show that large megathrust earthquakes seem to rupture the patches on the interface that are locked during the interseismic period and rarely propagate through creeping segments. The location and extent of such locked patches on plates interfaces therefore appears as a good proxy in order to anticipate the location and magnitude of megathrust earthquakes.

interseismic period, the slip on the fault at the interface between the two plates is heterogeneous: some parts of the interface creep steadily while some others are stuck. This so called “coupling pattern” can then be compared to the areas on the fault that slip during large events, and the coseismic slip pattern appears to be a negative of the interseismic slip on the interface, the areas of the fault that rupture during earthquakes correspond to the ones that were locked during the interseismic period before the event.

The first major component of my doctoral work, detailed in the publication in chapter 1, was therefore to determine the interseismic coupling pattern on the MHT under Nepal with as much accuracy as present day data allows, using surface deformation measured by the current Nepalese GPS network, which grew throughout the years from 3 stations around Kathmandu in 1997 to an extensive network of nowadays 30 stations covering all of Nepal.

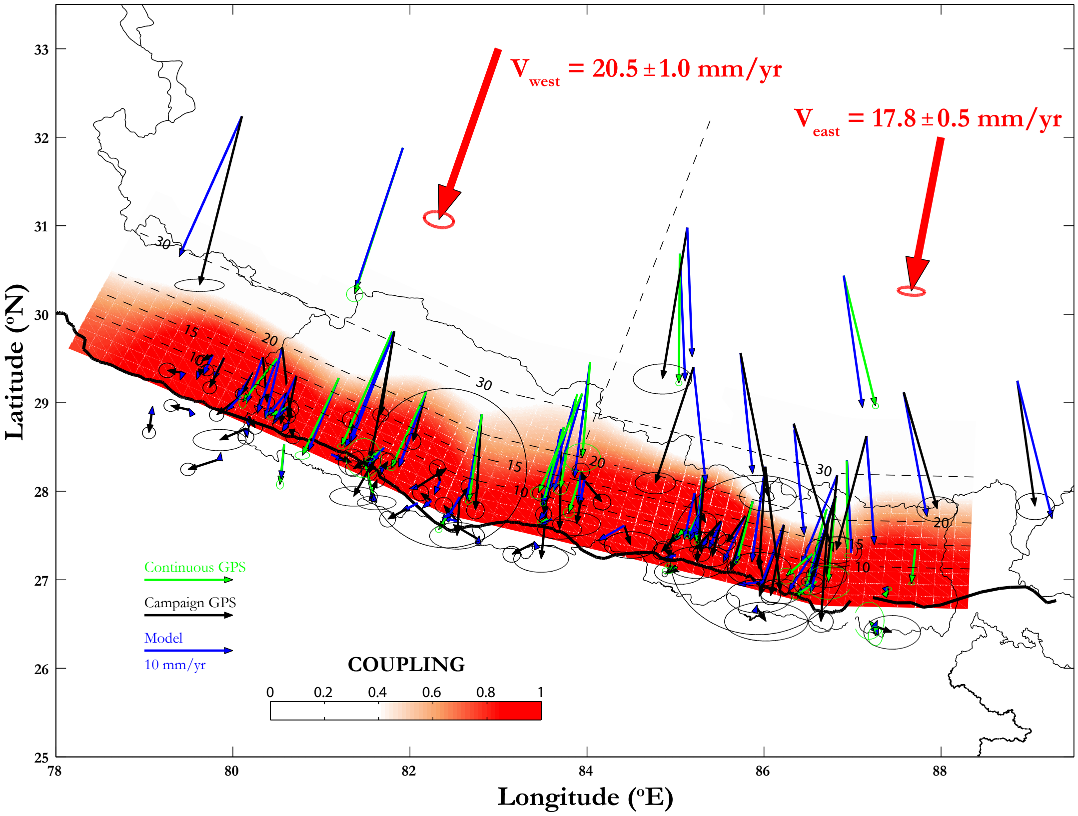


Figure 6: Comparison between observed horizontal GPS velocities and predicted by the coupling pattern. Interseismic coupling is shown as shades of red on the fault: red areas are locked during the interseismic loading, while white areas are creeping. The GPS data with corresponding error bars are plotted respectively as green and black arrows for the continuous and campaign GPS measurements. Blue arrows show predicted velocities according to the plotted pattern of interseismic coupling. Dashed line approximately trending north-south indicates the east-west separation, on each side of which the secular velocity can be different. Red arrows represent the east and west long term convergence rate across the Himalaya. Black dashed lines with numbers represent contour lines of fault depth (in km).

Figure 6 shows the coupling and interseismic convergence rate across the Nepal Himalaya, computed in the publication from chapter 1. Whereas the interseismic couplings on subduction zones interfaces always seem to exhibit patchy patterns (Figure 5), the locked patches representing potential locations for megathrust earthquakes to develop, the coupling pattern in Nepal is extremely homogeneous along strike: the MHT seems to be homogeneously locked from the surface to about 100 kilometers along dip, and then creeping at greater depths at an interseismic convergence rate of about 20 mm/yr. This interseismic convergence rate matches the long term slip rate at the front of the Main Himalayan Thrust of 21.5 ± 1.5 mm/yr, deduced by *Lavé and Avouac* (2000) from the uplift of Holocene terraces. This indicates that the interseismic deformation is mostly elastic, and that all the slip that occurs at depth during the interseismic period has to somehow propagate all the way to the front of the chain at some point, most likely during large earthquakes. These large earthquakes are rare, the last one in Nepal to have ruptured the MHT up to the surface being the 1934 $M_W \sim 8.4$ Bihar-Nepal earthquake (*Sapkota et al.*, 2013). No large event rupturing the MHT up to the surface has occurred during the instrumental period and has thus been recorded in Nepal, but the Himalaya produced a $M_W = 7.6$ earthquake in the Kashmir region in 2005 that ruptured the fault up to the surface and claimed 80,000 lives.

How large and where can these mega earthquakes be, are fundamental questions that we would like to be able to unambiguously answer. But where we were expecting to resolve isolated locked patches during the interseismic period, we see a homogeneous pattern unique to the Nepal Himalaya, which raises as many questions as it answers. The main one that now comes to mind is: can the entire fault rupture all at once in an unprecedented humongous earthquake?

From the point of view of the coupling pattern, there does not seem to be any creeping area that would act as a barrier against the propagation of large events. Similar studies in Bhutan and India suggest the same homogeneous along strike coupling pattern, and it thus seems that this homogeneous pattern extends to the whole Himalayan range. I also spent a fair amount of time during my PhD looking for repeating events within the belt of seismicity in order to obtain independent constraints for the convergence rate across the MHT, by correlating the waveforms of all seismic events recorded by the National Seismological Center (NSC) in Kathmandu, Nepal between 1995 and 2010. Given that the convergence rate is about 20 mm/yr and that the completeness magnitude of the seismic network is below $M_L = 3$, if some of the microseismicity on the MHT contained repeating events, they should be detected. But where other subduction zones usually display such events, less than 10 repeating earthquakes could be detected on the MHT, highlighting again the discrepancy between the MHT and other subduction zones.

Jumping to the conclusion that the Himalaya is able to produce seismic events that would rupture the entire arc all at once might however be a bit hasty. That a fairly circular patch locked in the interseismic period produces a large event does not necessarily mean that a rectangular homogeneously locked patch 3000 km long and only 100 km wide will necessarily be able to generate a single event, as geometrical factors will most likely come into play. Besides, unlike other subduction zones, the MHT seems to be locked all the way to the surface, and the effect of the free surface on a hypothetical along-strike propagation of a seismic rupture remains unclear.

Known Himalayan earthquakes that ruptured the locked part of the MHT all the way to the surface, such as the 2005 Kashmir or the 1934 Bihar-Nepal earthquakes, only ruptured a bounded along strike segment of the fault, which always remained short compared to the total length of the Himalayan range. But this has, unfortunately, no reason to mean that one day an event could not propagate throughout the entire arc.

The reasoning can be pushed further than these phenomenological considerations. Assuming that both this coupling pattern and the convergence rate remain steady in time, it is possible to quantify the rate of moment deficit, i.e., the moment that accumulates under Nepal every year, awaiting to be released somehow, possibly by large earthquakes. We played that game in *Ader et al.* (2012a) and computed a rate of moment deficit of $\dot{\mathcal{M}}_0 = 6.6 \pm 0.4 \times 10^{19}$ Nm/yr. This is a lot: this is a $M_W = 7.6$ event, similar to the 2005 Kashmir earthquake with its aftershocks, every 14 years. This is a magnitude 8.4 every 225 years.

However, one piece of the puzzle is missing in order to nail down such seismic hazard assessments: the magnitude of the largest possible event in the Himalaya, which is actually the information we were initially after. Magnitude 8.4 events would happen every 225 years in Nepal if these were the largest events ever produced on the MHT. But if one assumes that the MHT can produce events as large as the M_W 9.2 2004 Sumatra earthquake, then events of magnitude $M_W \geq 8.4$ would only need to happen every 570 years to balance the rate moment deficit.

Trying to deduce this information from the know Himalayan seismicity, Figure 7 plots the Gutenberg-Richter distribution of the Nepalese seismicity, i.e., the annual number of events above a given magnitude as a function of this magnitude, for three available catalogs: the catalog recorded by the National Seismological Center (NSC) in Kathmandu, Nepal between 1995 and 2001, the Centroid Moment Tensor (CMT) catalog covering the period from 1976 to 2010, and a historic catalog compiled by *Ambraseys and Douglas* (2004). These catalogs obviously have different sensitivities (i.e., different completeness magnitudes) but cover different time periods, therefore populating different regions of the Gutenberg-Richter plot.

All three distributions seem to line up along a line of slope of -1 (Gutenberg-Richter *b*-

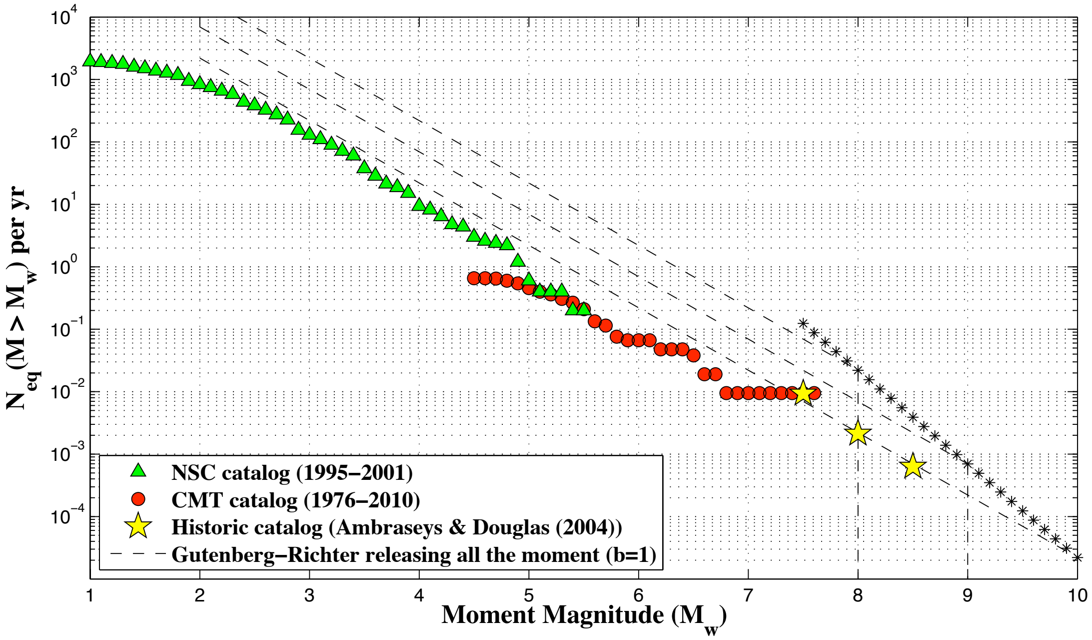


Figure 7: Gutenberg-Richter plot of the seismicity in Nepal, using the different catalogs available: The NSC catalog (1995-2001), the CMT catalog (1976 - 2010) and an historic catalog compiled using the catalog from *Ambraseys and Douglas* (2004). We used the last 500 years of the historic catalog for $M_W > 8$ earthquakes, and the last 200 years for $M_W > 7.5$ earthquakes. The dotted lines are the distribution that the seismicity should follow if 100% of the moment deficit was released seismically following a Gutenberg-Richter distribution with $b = 1$, up to a given maximum magnitude of 8, 9 and 10. The asterisk line shows, for a given maximum possible magnitude for Himalayan earthquakes, the return period of such earthquakes.

value of 1), and succumbing to the temptation of extrapolating these distributions to larger magnitudes in order to deduce the largest possible magnitude of events in Nepal would give $M_W^{\max} = 10$. Let's be clear, doing so would be simply and plainly wrong. This would be unreasonable first, because such $M_W = 10$ maximum magnitudes would require an average slip on the fault of a hundred meters, assuming a shear modulus of 30 GPa and a locked part of the MHT 3000 km long and 100 km wide. This would also be bluntly wrong because in order to be able to extrapolate the distribution to larger magnitudes, the seismicity plotted would have to be the average seismicity produced by the fault during a full seismic cycle. It would be possible only if the seismicity rate was constant through time, but we know that this is far from being the case. The seismicity rate increases drastically after large events, during aftershock sequences, meaning that the catalog duration would have to cover several complete seismic cycles for the seismicity considered to be a good estimate of that average, and therefore be longer than several times the return period of the largest possible event, i.e., several thousand of years. Even at magnitude 8.4, the 500 years covered by the historic catalog by *Ambraseys and Douglas* (2004) are too little time compared to the minimum possible return period of such events (225 years) to assure that the value plotted in Figure 7 is a good estimate of the average return period of such events.

As discussed in *Ader et al.* (2012a), the rate of moment deficit might vary with time, although such time variations probably have a minor impact on the global budget of moment deficit, as suggested by the fact that the coupling pattern is so homogeneous along strike, although a large event ruptured the MHT in eastern Nepal in 1934 while the western part of Nepal has not ruptured in the past 500 years. A fraction of this moment deficit might also be released in an aseismic way during the postseimic slip following large earthquakes or during slow slip events, but observations in other subduction zones suggest that this fraction would be small.

So the question is still open: how big can Himalayan earthquakes be? Now that we have explored our possibilities, it seems that the only remaining way to answer this question would be to numerically simulate the evolution of the MHT over a few tens or hundreds of thousands of years, i.e., over a few complete seismic cycles, and see if it is possible to both produce events that rupture a finite segment of a uniformly locked MHT up to the surface and occasionally events which rupture the entire arc. The hurdles standing in the way of realizing such a simulation are still manifold, and some of them are still out of our reach, but some issues can already be tackled with the tools we have in hand.

We already know fairly well the geometry of the locked zone and the convergence rate thanks to GPS data, but we need to determine the physical parameters characterizing the fault properties that we will plug into such a simulation. Friction between rocks and therefore fault behavior is usually described with the rate-and-state framework (*Dieterich*, 1979a,b; *Ruina*, 1983).

In the rate-and-state formalism, the evolution of the friction coefficient μ between two rock surfaces or gouge layers logarithmically depends on the slip rate V and a state variable θ (*Dieterich*, 1978, 1979a,b; *Ruina*, 1983):

$$\mu = \mu^* + a \ln \frac{V}{V^*} + b \ln \frac{\theta V^*}{D_c}, \tag{1}$$

where μ^* is the reference friction coefficient corresponding to the reference slip velocity V^*, D_c is the characteristic slip for state evolution (e.g., *Dieterich*, 1978, 1979a,b; *Ruina*, 1983; *Rice and Ruina*, 1983; *Dieterich and Kilgore*, 1994), and a and b are rate-and-state constitutive fault parameters. The state variable θ can be interpreted as the average age of the population of contacts between two surfaces and its evolution is usually described either by the aging law (e.g., *Marone*, 1998):

$$\frac{d\theta}{dt} = 1 - \frac{V\theta}{D_c}, \tag{2}$$

or by the slip law (e.g., *Marone*, 1998):

$$\frac{d\theta}{dt} = -\frac{V\theta}{D_c} \ln\left(\frac{V\theta}{D_c}\right). \tag{3}$$

Note that the state variable evolves in time even if there is no relative motion between rocks in contact.

In the steady-state regime where $d\theta/dt = 0$, $V_{\text{ss}} = D_c/\theta$ and the friction can therefore be expressed as a function of the slip rate on the fault only:

$$\mu = \mu_{\text{ss}} + (a - b) \ln \frac{V}{V_{\text{ss}}}. \tag{4}$$

Equation (4) shows that the friction has two distinct behaviors depending on whether $a-b$ is positive or negative. When $a-b$ is positive, the friction on the fault increases when the slip rate increases, which promotes stable creep (rate-strengthening rheology). Conversely, in the case where $a - b$ is negative, the friction decreases when the slip rate increases, which might lead to highly increasing slip rates (rate-weakening rheology). Looking at the coupling pattern in Nepal in Figure 6, one can easily see how this formalism can qualitatively represent the observed pattern on the fault: the creeping part of the fault may be described by a rate-strengthening rheology, while the locked zone would be represented by a rate-weakening rheology. Quantitatively though, in order to run a realistic simulation, the numerical values of rate-and-state fault parameters a, b and D_c would need to be determined.

It turns out that fault properties may be extracted by analyzing the response of the fault to stress perturbations, and this is where Nepal represents an incredible natural laboratory with a helpful variety of stress perturbations. On top of the secular shear stress loading at the plate interface resulting in a fairly constant stress rate on the MHT, two types of additional stress perturbations are at play. The first ones, perhaps the most common form of stress perturbation in a seismic zone, are the sudden stress changes due to large events. The study of the evolution of aftershock sequences may thus be able to shed light onto some of the local fault properties.

The second ones are periodically varying stresses. These are due to the solid Earth tides, which induce variations of the shear stress on the MHT of the order of 3 kPa of amplitude (*Bettinelli et al.*, 2008), or, a specificity of the Himalaya, to the hydrological cycle dominated by the heavy monsoonal rains in the summer, which causes a dramatic accumulation of water in the aquifers of the Gangetic basin. This surface load bends the Indian plate, generating annual shear stress variations on the MHT between the wet summer and the dry winter months, also of about 3 kPa of amplitude. The individual and relative responses of the fault to both these periodic stress variations are likely constrained by fault properties, and the second major component of my PhD has thus been to evaluate the amplitude of these responses in Nepal.

The Nepalese microseismicity from 1995 to 2008 is plotted in Figure 8, both in map view and in terms of the evolution of the cumulative number of events with time. In order to appraise the response of this seismicity to both periodic stress perturbations (seasonal variations and tidal perturbation), I developed a tool based on the Schuster test to assess the existence of periodicities in the timing of events in an earthquake catalog, which is presented in the manuscript submitted for publication in chapter 2. With this tool, called the Schuster spectrum, I pointed out that looking at small magnitude events in a catalog may provide a statistically more significant number of events, but these events quickly tend to cluster in time (aftershock sequences, clusters, etc.) and any periodicity in the catalog may be occulted by this loss of independence between events. The best way to circumvent this problem is hence to look at the largest events in a catalog which will most likely remain uncorrelated with each other. As a vicious consequence, looking at a region over a longer time period will not necessarily increase the number of events available to assess the existence of any periodicity: the magnitudes of the largest events present in the catalog increasing with time, the magnitude of events that one should use will also increase.

In order to examine the periodicities in the seismicity on the MHT we thus compute the Schuster spectrum (*Ader and Avouac*, 2013) for $M_L \geq 5.5$ events from the National Seismological Center (NSC) seismic catalog in Nepal from 1995 to 2008, and $M_b \geq 4$ events from the International Seismological Centre (ISC) catalog from 1965 to 2008. Both spectra are reproduced from *Ader and Avouac* (2013) in Figure 9 and show an annual variation of the seismicity rate in Nepal above the 95% confidence level, but no detectable variation at any of the tide periods. This indicates that even though the seasonal and tidal stress perturbations have comparable amplitudes, the response of the seismicity is larger in the

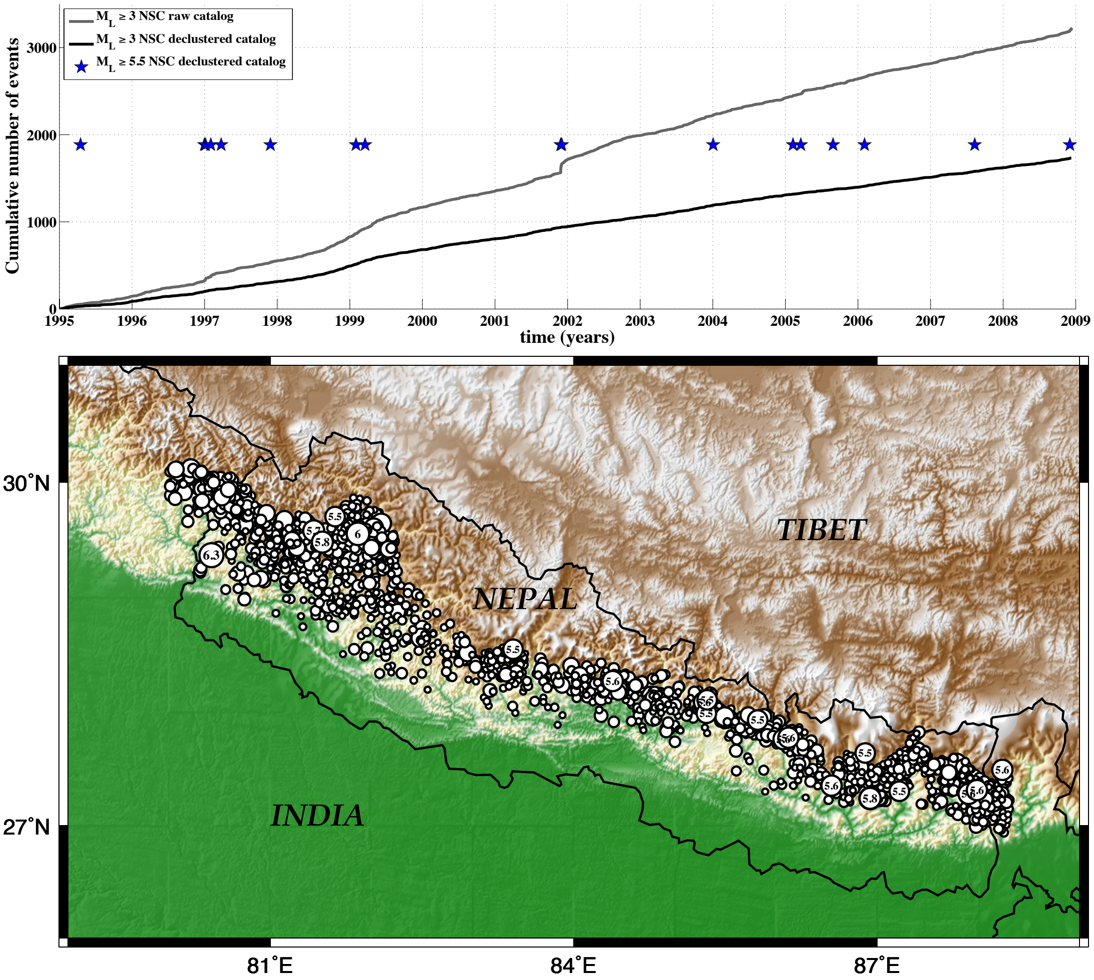


Figure 8: Time and space distribution of the $M_L \geq 3$ NSC seismicity used in this study. Upper plot shows cumulative number of events from 1995 to the end of 2008 for raw (grey curve) and declustered (black curve) catalogs, together with times of $M_L \geq 5.5$ events from the declustered catalog (blue stars). The map shows midcrustal events from the raw catalog used in the study, selected according to their localization, using the same selection contour as in *Bollinger et al.* (2007). Circles sizes are proportional to events magnitudes: smallest events have $M_L = 3$, and $M_L \geq 5.5$ events are indicated by their magnitude, giving an idea of the scale.

annual case. Quantitatively, we show that the relative amplitude of variations of seismicity rate could be as large as 40% at the annual period, while the amplitude of the response at the tidal periods is less than 25%.

As is represented in Figure 6, the MHT is made of two distinct zones: a creeping zone at depth, best modeled by a rate-strengthening rheology in the rate-and-state framework, and a locked seismogenic zone, that would be described by a rate-weakening rheology. I

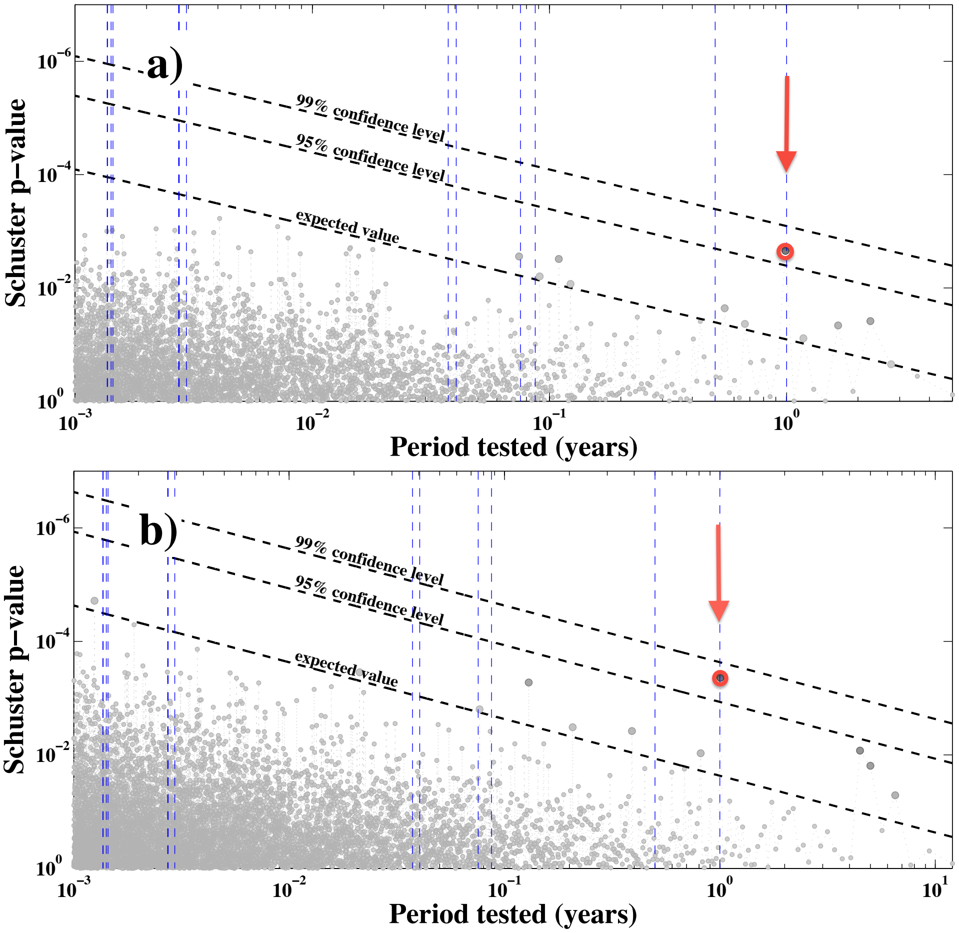


Figure 9: Schuster spectra (*Ader and Avouac*, 2013) computed for large events from a) the NSC and b) the ISC declustered catalogs. For the NSC catalog, the spectrum is computed for $M_L \geq 5.5$ events for the entire available catalog (1995 to 2008). The spectrum for the ISC catalog is computed for $M_b \geq 4$ events from 1965 to 2008. The periodicity at one year is highlighted in red and indicated by a red arrow.

therefore studied the response of both rheologies to harmonic shear stress variations, in order to evaluate the impact that both zones may have on the response of seismicity to both tidal and seasonal stress perturbations.

First and foremost, I started with the response of a rate-strengthening fault to a harmonic variation of Coulomb stress, which is detailed in the publication in chapter 3. The Coulomb stress $S(t)$ is defined in terms of the shear stress $\tau(t)$ and the normal stress $\sigma(t)$:

$$S(t) = \tau(t) - \mu\sigma(t), \tag{5}$$

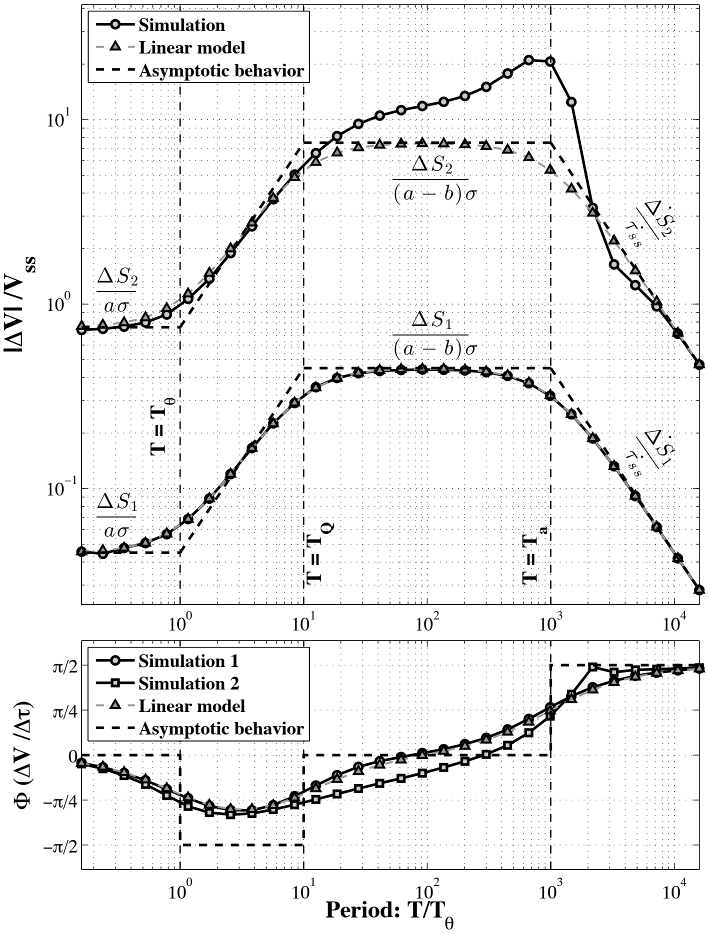


Figure 10: Response of a spring-slider system with rate-strengthening rheology to small harmonic Coulomb stress perturbations of different periods and amplitudes $\Delta S_1 = 0.9$ kPa (simulation 1) and $\Delta S_2 = 15$ kPa (simulation 2). The amplitudes of the shear and normal stress perturbations are the same. The system is undergoing constant loading at velocity $V_{\text{ss}} = 0.02$ m/yr under mean normal stress $\sigma_o = 5$ MPa. The normalized spring stiffness is $k/\sigma_o = 0.002$ m^{-1}. The other parameters are: $\mu_o = 0.7$, $a = 0.004$, $b = 0.0036$ and $D_c = 2\times10^{-4}$ m. Upper panel: Amplitude of the creep rate variations. The black lines with circles represents the results of the simulations (one line for each value of ΔS). The dashed grey lines with triangles represent the small perturbation approximation (*Ader et al.*, 2012b) for each simulation while the dashed light grey lines indicate the corresponding asymptotic behavior of the system with equations indicated on the plot. The critical periods T_θ, T_Q and T_a are also indicated on the plot. Lower panel: Phase difference between the creep rate variations and the stress perturbation.

where μ is the friction coefficient. Creeping faults do not produce large earthquakes since no moment deficit is usually accumulated on them, but they seem able to produce tremor like signals, which intensity is usually considered to vary with the creep rate on the fault (*Ide et al.*, 2007; *Shelly et al.*, 2011). Tremor intensity and therefore slip rate on creeping faults has been reported on multiple occasions to highly correlate with tidal perturbations (*Rubinstein et al.*, 2008; *Nakata et al.*, 2008; *Thomas et al.*, 2009, 2012; *Hawthorne and Rubin*, 2010) in spite of the low amplitude of tidal perturbing stresses, indicating a high sensitivity of creeping faults to stress perturbations. This high sensitivity is generally attributed to near lithostatic pore pressures, which would reduce the effective normal stresses by orders of magnitude below their lithostatic value at the depths considered.

As has been highlighted in the case of non-volcanic tremors (*Ader et al.*, 2012b), near-lithostatic pore pressures require specific fault properties for the nucleation sizes to remain consistent with the occurrence of the smallest earthquakes recorded. Besides, near-lithostatic pore pressure and corresponding effective normal stresses are orders of magnitude below the values reported from afterslip studies in various tectonic contexts (*Hearn et al.*, 2002; *Miyazaki et al.*, 2004; *Perfettini and Avouac*, 2004, 2007; *Hsu et al.*, 2006, 2009a,b; *Fukuda et al.*, 2009; *Barbot et al.*, 2009).

The main point made in *Ader et al.* (2012b), presented in chapter 3, is that a set of rate-and-state parameters such that $a - b \approx 0$ on a creeping fault can cause a highly amplified response of the creep rate to a harmonic stress perturbation. At the right perturbing period, the amplitude of the variations of slip rate due to a perturbation of Coulomb stress of amplitude ΔS becomes indeed

$$\frac{\Delta V}{V_{\text{ss}}} \approx e^{\frac{\Delta S}{(a-b)\sigma}}, \tag{6}$$

which can be large if either $a - b$ or σ are small enough for $(a - b)\sigma$ to be smaller than the amplitude of the stress perturbation.

Figure 10, taken from *Ader et al.* (2012b), shows the response of a spring-slider system with rate-strengthening rheology to harmonic perturbations of Coulomb stress throughout a range of periods, for two different amplitudes ΔS_1 and ΔS_2 such that $\Delta S_1 < (a - b)\sigma$ and $\Delta S_2 > (a - b)\sigma$. It shows that between the characteristic periods T_Q and T_a (see *Ader et al.* (2012b) for the expressions and physical meaning of these periods), the amplitude of the response depends on $(a - b)^{-1}$. Within this range of periods, when $\Delta S \ll (a - b)\sigma$, equation (6) can be linearized to $\Delta V/V_{\text{ss}} \approx \Delta S/(a - b)\sigma$. But when $\Delta S_2 > (a - b)\sigma$ the amplitude of the response becomes larger than this linear approximation: the response has a large amplitude and is described by equation (6), as is derived and tested in *Ader et al.* (2012b). On a fault region where $a-b \approx 0$, this second regime is consequently likely to be at play. This especially applies to slow slip events and tremors since they are usually observed at the transition between the locked and creeping parts of faults, i.e., as was mentioned earlier, at the transition between a rate-weakening ($a-b<0$) and and a rate-strengthening ($a-b>0$) rheologies. The delimitating periods T_a, T_Q and T_θ in Figure 10 are inversely proportional to the background loading rate, and thus the sensitivity to periodic stresses should vary when this background velocity changes, as for example during a slow slip event. Such variations of response amplitude have been observed in Parkfield, California (*Thomas et al.*, 2012), and might therefore help bring constraints on fault parameters on the San Andreas fault.

Tremors and slow slip events yet remain to be observed in Nepal, and if they exist there, the analysis of their response to tidal and seasonal stress variations should be able to bring tight constraints on the fault parameters of the creeping part of the MHT. It is not yet clear if annual variations of the creep rate are responsible for the observed variations of the seismicity rate on the MHT. The variations of strain recorded at the surface by the GPS stations in Nepal seem to be entirely explained by the elastic deformation of the crust in response to the hydrological surface loading, and do not display any significant signal that would come from variations of creep on the MHT. If small variations of the creep rate at depth might not be large enough to generate detectable strain variations at the surface, there is still the possibility that they induce important variations of stress at the transition between the locked part and the creeping part of the MHT, cradle of the Himalayan seismicity. Such effects still need to be studied and quantified.

I then focused my efforts on the locked seismogenic part of the MHT in order to determine if a fault with rate-weakening rheology could explain both the annual variations of seismicity in Nepal and the apparent lack of response to tidal perturbations. Trying to explain this contrasted response of the seismicity with a simple 1D Spring-slider Rate-and-state Model (hereafter referred to as SRM) as we did earlier in the case of a rate-strengthening fault proves to be unsuccessful, as a spring-slider with a rate-weakening rheology would predict a response to the tides at least as large as for the annual forcing, as is showed in Figure 11 (*Ader et al.*, 2013). This is due to the fact that if a simple spring-slider is appropriate to describe the behavior of a creeping fault, which almost always remain in its steady-state regime, it does not reproduce all the complexity of the transitory earthquake

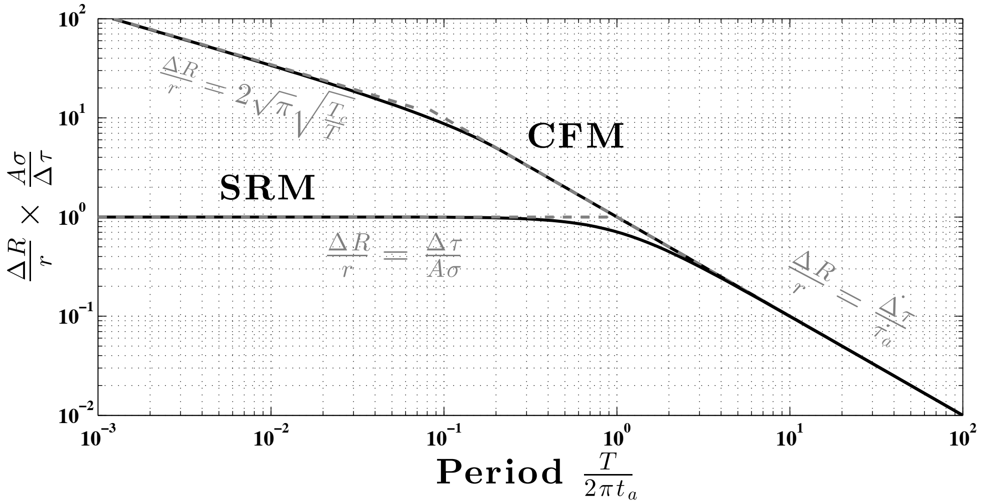


Figure 11: Amplitude of seismicity rate variations on the fault for different periods of shear stress harmonic variations of constant amplitude, according to the Coulomb Failure model (CFM) and the SRM. The predicted curve of the SRM is derived from the equations presented in *Dieterich* (1994). See *Ader et al.* (2013) for details on both models. The black curves represent the exact solutions solving for equations numerically, while the grey dashed lines represent the asymptotic behavior at large and short perturbing periods, with corresponding equations indicated on the plot, and derived in *Ader et al.* (2013). Supposing that the response to the Monsoon is on one of these curves, neither the SRM nor the Coulomb models are able to explain the fact that the response of Nepalese seismicity to tides is of less amplitude than the response to the Monsoon, remembering that both perturbations have the same amplitude, and should thus be one the same curve, but that the tides should be to the left of the monsoon on this plot as they have a much smaller period.

generation processes, and therefore leads to wrong quantitative predictions in the case of a rate-weakening rheology. Indeed, the generation of an earthquake on a rate-weakening patch starts with the nucleation phase, where a creeping zone at the edge of the patch, called the nucleation zone, steadily grows in size under the influence of the remote secular loading. When the nucleation zone reaches a critical size, it becomes unstable and starts growing recklessly: this is the propagation phase of the seismic rupture, or simply the seismic event itself. If the notion of a critical nucleation size appears in the spring-slider model, the notion of a physically growing nucleation zone does not exist.

A major task of my doctoral work has thus been to study the evolution of rate-weakening patches under harmonic stress perturbations, resorting to 2D fault models, which constitutes the last chapter of my PhD thesis. This problem is too complex to be entirely studied analytically and the evolution of slip on a finite fault was thus numerically simulated using the BICYCLE (Boundary Integral Cycles of Earthquakes) code developed at Caltech, based on the boundary integral method, and described in *Lapusta et al.* (2000), *Lapusta and Rice* (2003) and *Lapusta and Liu* (2009). The actual code used in the simulations presented in *Ader et al.* (2013) is the one presented in *Noda and Lapusta* (2010), where we have added the possibility of a perturbation of stress history superimposed over the entire fault. The results are presented in a publication in preparation for the Journal of Geophysical Research reproduced in chapter 4.

Figure 12 gives a schematic representation of the fault for which we compute the evolution through time. BICYCLE uses a spectral representation of the boundary integral formulation (*Lapusta et al.*, 2000) in order to deal with the computationally intensive inertial effects, which requires to periodically pave an infinite 2D space with the fault from Figure 12, resulting in a fault infinite in the direction indicated as a dashed line in Figure 12. It is made of a rate-weakening seismogenic patch (dark blue in Figure 12) embedded within a rate-strengthening creeping matrix (yellow part of the fault). The fault is only 3 km wide, with the seismogenic patch being 500 m in most of the simulations. The cell size used depended on the physical parameters of the simulation, but was 0.5 m in the vast majority of the simulations and for all figures that are presented in this introduction. The secular loading on the fault is reproduced by loading the fault at its edges with a constant velocity V_{pl}. When imposed, the stress perturbation is applied over the entire fault.

The second main ingredient of BICYCLE is the variable time stepping (*Lapusta et al.*, 2000; *Lapusta and Rice*, 2003), which enables the simulation of both the quiet interseismic period of quasi-static deformation with large time steps but also the rapidly evolving

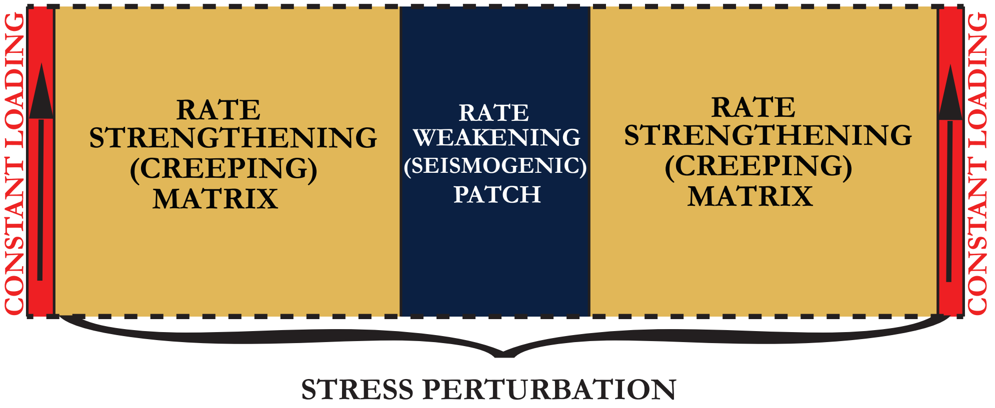


Figure 12: Cartoon showing the finite fault used in the numerical simulations. The fault is infinite in one direction, and consists of one rate-weakening (seismogenic) patch embedded within a rate-strengthening medium. The whole fault is loaded on both side at a constant velocity. The stress perturbation (either a step function or a harmonic perturbation) is applied over the entire fault. In most simulations in *Ader et al.* (2013), the fault has a length of 3 km, and the seismogenic patch at the center is 500 m long.

nucleation and propagation processes during seismic events. In order to simulate a statistically significant number of seismic events and seismic cycles, this algorithm was run during ten days on two HP SL390 Compute Nodes, each containing twelve Dual Westmere X5650 processors running at 2.67 GHz. On an fault evolving under the sole influence of the secular loading at the edge and without any additional stress perturbation, this produced a seismic catalog of about 15,000 events spanning a period of over 1700 years. Figure 13 shows approximatively two year of the slip evolution on the fault, and displays the diversity of seismic events and cycles produced by such a fault. This Figure also indicates that the duration of a seismic cycle, i.e., the time between two major events rupturing the entire seismogenic patch, never exceeds half a year, so that a 1700 year-long seismic catalog covers a large number of complete seismic cycles.

The fact that the natural evolution of slip on the fault (i.e., under the sole influence of the loading at the edges of the fault) and population of seismic events produced by such a finite fault displays some complexity is already a notable difference from the periodic uniform seismicity predicted by a simple spring-slider model. Since the fault is infinite in one direction, the magnitude considered here is the magnitude per unit of length in the infinite direction, and is defined as:

$$M_{\text{lin}} = \frac{2}{3}\log_{10}\mathcal{M}_{\text{lin}} - 6.7, \tag{7}$$

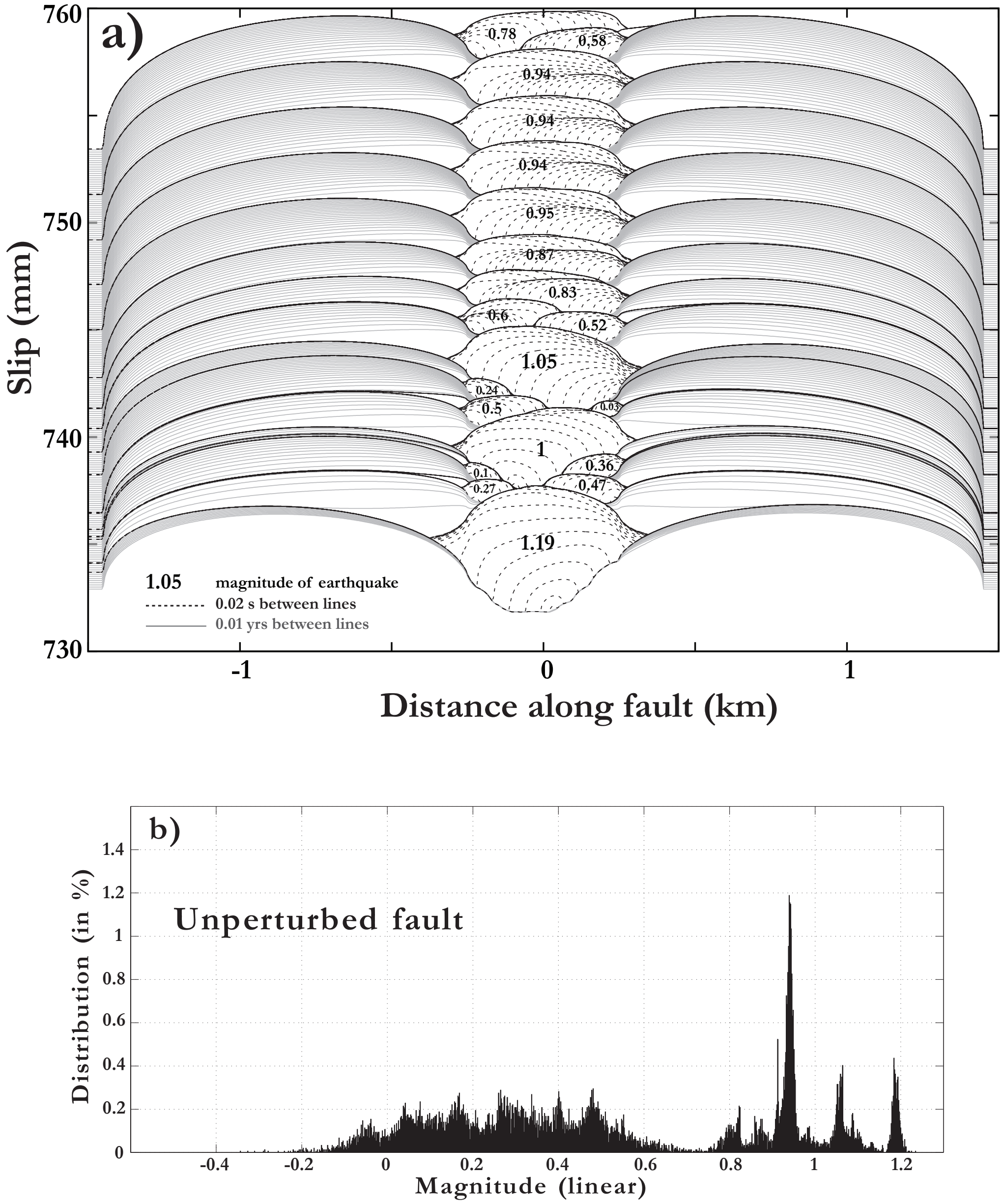


Figure 13: Upper plot: Evolution of slip during about 2 years on the unperturbed fault loaded at the edges at a constant velocity $V_{pl} = 1$ cm/yr. The linear magnitude of the seismic events produced is indicated on each event. If the fault is in the coseismic regime, the slip on the fault is plotted every 0.02 s (dashed black lines), while during the inter seismic period, slip is only plotted every 0.01 yrs (plain grey lines). Lower plot: Magnitude distribution of events produced by the fault. The earthquake catalog produced by the simulation contains a total of about 15,000 events, covering about 1700 years of evolution of the fault.

where the linear moment $\mathcal{M}_{\text{lin}}$ of a seismic event on this fault is:

$$\mathcal{M}_{\text{lin}} = G \int_{\text{fault}} s(x)\, dx, \tag{8}$$

with $x \in [-1.5; 1.5]$ km the position along the fault, $s(x)$ the slip on the fault at position x and $G = 30$ GPa the shear modulus. The magnitudes of earthquakes produced in the upper plot of Figure 13 are indicated in the plot and give an insight on which type of event the magnitudes in the lower plot correspond to: the bimodal distribution of magnitudes corresponds to events that either rupture the whole seismogenic patch ($M_{\text{lin}} > 0.75$) or only rupture an edge of the patch ($M_{\text{lin}} < 0.75$).

Response of a finite rate-and-state fault to harmonic shear-stress perturbations. Each point on the plot corresponds to a simulation which generated an earthquake catalog. The times of events from the catalog are stacked over one period and the resulting stacked seismicity rate is fitted with equation (4.19). The value obtained for β is reported on the upper plot, while the phase shift Φ between the seismicity rate and the stress is represented on the lower plot. Simulation parameters are described in *Ader et al.* (2013): $a = 0.008$, $b = 0.004$ in the creeping zone and $b = 0.012$ on the seismogenic patch. Normal stress is $\sigma = 5$ MPa and parameter $D_c = 5$ μm, and the loading velocity is $V_{pl} = 1$ cm/yr. Dashed grey curves show the predictions of the SRM, where we have taken $2\pi t_a = 0.1$ years in order to fit the phase. As in the SRM, one can separate two regimes of response, depending on whether the perturbing period is shorter or greater than a critical period T_a. However, the amplitude of the response is always greater than predictions from the SRM, sometimes by more than an order of magnitude, and this amplitude of the response increases with the period T for periods $T < T_a$.

This fault is then perturbed in *Ader et al.* (2013) with a harmonic shear stress perturbation and we look at the response of the seismicity. Figure 14 displays the harmonic response of the finite fault from the simulations, and is thus the equivalent of Figure 11 for a finite fault. The response for the SRM is reminded as a dashed grey line in Figure 14.

The way Figure 14 is obtained is detailed in *Ader et al.* (2013), but I summarize here the pertaining points. Each grey circle corresponds to one simulation, where the finite fault undergoes a harmonic shear stress perturbation of period T, indicated on the x-axis, and of amplitude $\Delta\tau$, constant from one simulation to another. In order to evaluate the variations of seismicity rate with time in the output seismicity catalog, we stack the times of events from the catalog over one period T and fit the obtained seismicity rate over one period with

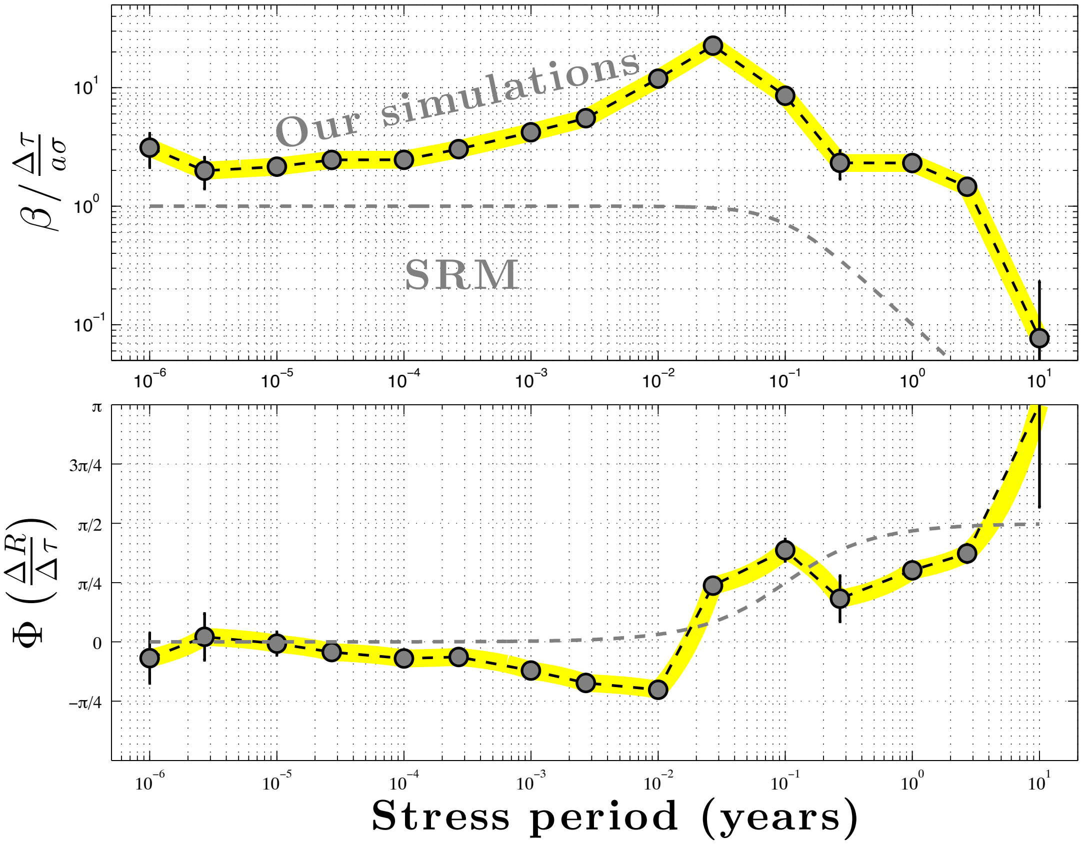


Figure 14: Response of a finite rate-and-state fault to harmonic shear-stress perturbations. Each point on the plot corresponds to a simulation which generated an earthquake catalog. The times of events from the catalog are stacked over one period and the resulting stacked seismicity rate is fitted with equation (9). The value obtained for β is reported on the upper plot, while the phase shift Φ between the seismicity rate and the stress is represented on the lower plot. Simulation parameters are described in *Ader et al.* (2013): $a = 0.008$, $b = 0.004$ in the creeping zone and $b = 0.012$ on the seismogenic patch. Normal stress is $\sigma = 5$ MPa and parameter $D_c = 5$ μm, and the loading velocity is $V_{pl} = 1$ cm/yr. Dashed grey curves show the predictions of the SRM, where we have taken $2\pi t_a = 0.1$ years in order to fit the phase. As in the SRM, one can separate two regimes of response, depending on whether the perturbing period is shorter or greater than a critical period T_a. However, the amplitude of the response is always greater than predictions from the SRM, sometimes by more than an order of magnitude, and this amplitude of the response increases with the period T for periods $T < T_a$.

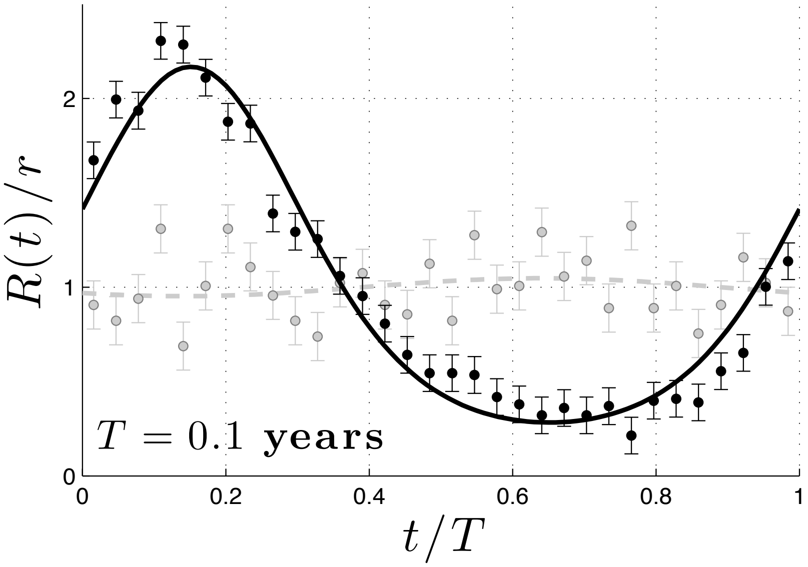


Figure 15: Stacked seismicity rate on a finite fault under a harmonic stress perturbation of period $T = 0.1$ years and amplitude $\Delta\tau = 3$ kPa. The black dots with error bars represent the normalized stacked seismicity rate from the perturbed catalog, computed by dividing the period into $B = 32$ bins of equal length and counting the number of events falling within each bin. The black line shows the fit of equation (9) with best fitting parameters: amplitude $\beta = 1.02 \pm 0.04$ and phase $\Phi = 35^\circ \pm 3^\circ$. The error bars on the seismicity rate only depend on the total number of events N in the catalog and the number of bins used to compute the seismicity rate from the seismicity catalog output from the numerical simulation: $\sigma_B^2 = (B-1)/N$ (e.g., *Ader and Avouac*, 2013). For comparison, the light grey circles with error bars show the seismicity rate from the unperturbed catalog stacked over the same period with corresponding fit of equation (9) with dashed grey line. In spite of being derived with the spring-slider model, equation (9) provides a good qualitative fit to the seismicity rate variations on the finite fault. However, given that the normal stress on the finite fault is $\sigma = 5$ MPa, and that the fault parameter $a = 0.008$, the spring-slider model would prescribe $\beta_{\mathrm{SRM}} = \Delta\tau/a\sigma = 0.075$, a value more than an order of magnitude less than the β-value required to fit the seismicity rate.

the following expression:

$$\frac{R(t)}{r} = \frac{e^{\beta \sin(\omega t - \Phi)}}{\langle e^{\beta \sin \omega t} \rangle}, \tag{9}$$

where the amplitude β and phase Φ of the response are determined to fit the results of each simulation. The notation $\langle . \rangle$ refers to the mean of the function. Equation (9) corresponds to the predictions of the seismicity rate by the SRM (this expression is derived in *Ader et al.* (2013) on the ground of the equations obtained by *Dieterich* (1994)), except that in the SRM, parameters β and Φ are prescribed in terms of the fault parameters.

A lot can be learned from these simulations and numerous results are highlighted in the paper in preparation reproduced in chapter 4. In order to explain the contrasted response of the seismicity in Nepal to tidal and seasonal stress variations on the fault, some of these conclusions come handy. Figure 15 may suggest that although predicted by the SRM, equation (9) seems to fit well the seismicity rate of the finite fault. However, quantitatively looking at the amplitude of the variations of the seismicity rate on the finite fault, the best fit is obtained for $\beta = 1.02 \pm 0.04$ whereas the SRM would advocate $\beta_{\text{SRM}} = \Delta\tau / a\sigma = 0.075$, a value more than an order of magnitude less than the β-value required to fit the seismicity rate. Reciprocally, interpreting this observed $\beta \approx 1$ amplitude on the finite fault with the predictions of the SRM would lead to $a\sigma \approx \Delta\tau = 3$ kPa, whereas the actual value of $a\sigma$ on the finite fault is 40 kPa. This underestimation of the amplitude of the response of seismicity to harmonic stress is not specific to the period $T = 0.1$ year, as Figure 14 shows that the amplitude of the variations of seismicity rate predicted by the SRM is systematically much lower than the amplitude yielded by the finite fault simulations. Estimating fault parameters on real faults by fitting variations of the seismicity rate observed in nature with predictions of the SRM (*Cochran et al.*, 2004; *Bettinelli et al.*, 2008) will therefore deceivingly yield a seemingly good fit, but as highlighted in *Ader et al.* (2013), the quantitative estimates of fault parameters may be off by sometimes several orders of magnitudes. The claim made by *Bettinelli et al.* (2008) that $a\sigma$ has to be as low as 8 kPa on the MHT (requiring either values of the fault parameter a orders of magnitudes lower than laboratory measured values (*Dieterich*, 1994) or near lithostatic pore pressure) for the microseismicity to responds as much as it does to seasonal variations of stress might therefore not be necessary. New laws relating rate-and-state fault parameters to the quantitative response of a rate-and-state fault to a harmonic stress perturbation have to be established in order to pull out trustworthy estimates of fault parameters from observations.

As for the fact that the seismicity on the MHT seems to have a higher response at the annual than at the tidal periods (*Ader and Avouac*, 2013) although both stress variations have comparable amplitudes (*Bettinelli et al.*, 2008), Figure 14 seems to indicate that for periods smaller than a critical period T_a, the amplitude of the response increases with the period, and then decreases as the period keeps increasing above T_a. The phase Φ goes from about 0 to about $-\pi/4$ as T grows towards T_a, indicating a seismicity rate mostly in phase with the shear stress with a slight time lag appearing as T grows closer to T_a, and then jumps to $\Phi \sim \pi/2$ when $T > T_a$, meaning that the seismicity rate becomes in phase with the shear stress rate. Applying these simple observations to the case of Nepal, the fact that the annual response of the seismicity is in phase with the stress rate but that it is still larger than the response to tides indicates that T_a must be of the order of 1 year. If T_a was much larger than 1 year, then the seismicity rate in Nepal would be in phase with the seasonal stresses and not the stress rates, but if T_a was much smaller than 1 year, then the amplitude of the response at 1 year would be very small and unlikely to be greater than the response to tides.

This value of T_a for Nepal can also be independently obtained by looking at aftershock sequences. We simulated in *Ader et al.* (2013) the response of a finite fault to a step of shear stress of amplitude $\Delta\tau$, and showed that the evolution with time of the cumulative number of event following the step (supposedly imposed at time $t = 0$) could be fitted by an expression again inspired by the predictions of the SRM:

$$N(t) = rt + r_a t_a \ln\left[e^{\beta} + \left(1 - e^{\beta}\right) e^{-t/t_a}\right] \mathcal{H}(t), \tag{10}$$

where the background seismicity rates r and r_a are free parameters to fit for, r being the background seismicity rate of the entire zone considered and r_a corresponding to the one only on the faults producing aftershocks. In the SRM, since only one fault is considered, both background rates r and r_a are equal. But looking at natural seismicity, one needs to consider a zone that encompasses the entire aftershock sequence, which will therefore also contain faults not responding to the stress change, which justifies resorting to two different background seismicity rates in equation (10). The two other free parameters are the amplitude of the response β and the characteristic relaxation time of the seismicity rate following the stress step (i.e., the characteristic duration of the aftershock sequence) t_a. The function $\mathcal{H}(t)$ is the Heavyside function: $\mathcal{H}(t) = 0$ for $t < 0$ and $\mathcal{H}(t) = 1$ for $t \geq 0$. Expressions of t_a and β are prescribed in the SRM in terms of other parameters of the problem: noting $\dot{\tau_a} = kV_{pl}$ the secular loading rate of shear stress on the slider,

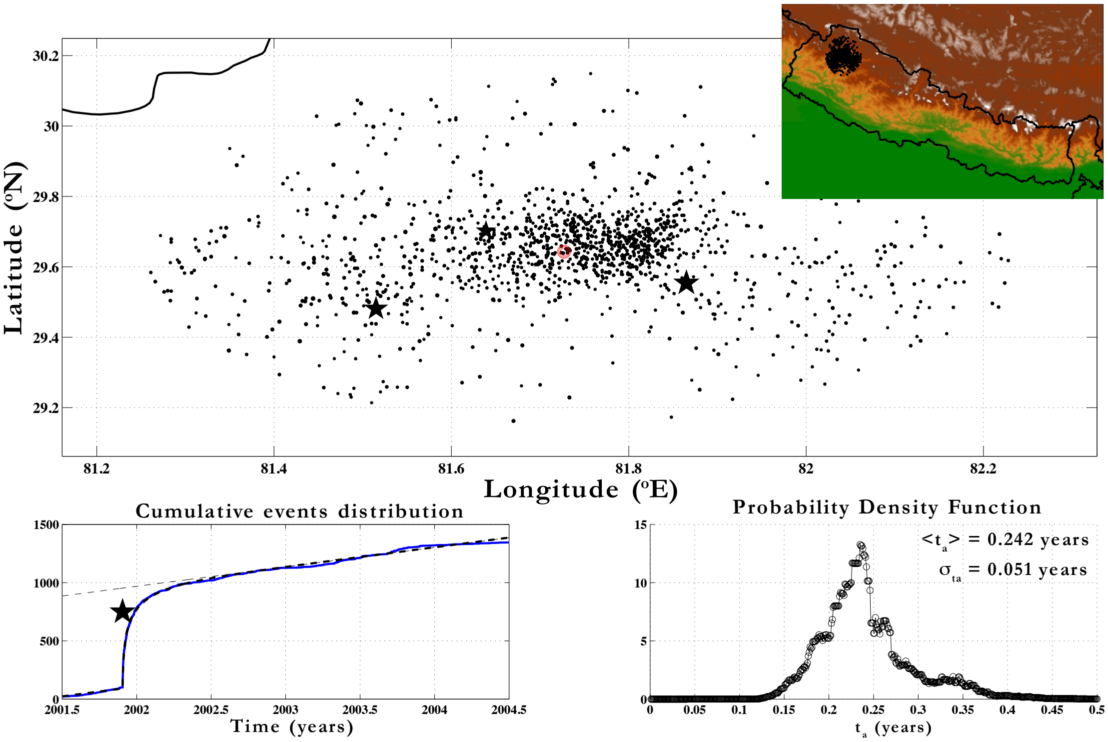


Figure 16: Aftershock sequence in Western Nepal following the $M_L = 6.1$ event that happened on November 27, 2001. The upper maps show the location of all $M_L \geq 2$ aftershocks within 0.5° from the center of the cluster, $M_L \geq 5$ events being signaled by a star. Lower left plot: cumulative number of events within this radius and fit with equation (10) for the best values of β and t_a. Best fit model has $t_a \approx 0.24$ years. Lower right plot: complete PDF for t_a, indicating that $t_a \approx 0.24 \pm 0.05$ years. This value is of the same order of magnitude as critical period $T_a \approx 1$ year, estimated from the frequency response of the seismicity in Nepal.

the characteristic relaxation time is $t_a = a\sigma/\dot{\tau}_a$, and the amplitude is $\beta_{\mathrm{SRM}} = \Delta\tau/a\sigma$, equivalently to the response to a harmonic perturbation of shear stress. We show in *Ader et al.* (2013) that similarly to the findings of *Kaneko and Lapusta* (2008) and to the harmonic response, the β-value necessary in order to fit the cumulative number of events produced by the finite fault in response to the step of shear stress is much larger than β_{SRM} and is thus hard to interpret in terms of the fault constitutive parameters.

In the SRM, the characteristic time t_a and the characteristic period T_a are related by the simple equation:

$$T_a = 2\pi t_a. \tag{11}$$

Assuming that this expression still remains valid in the context of a real fault, we compute t_a for the aftershock sequence following the $M_L = 6.1$ earthquake in western

Nepal on November 27, 2001, which is the largest aftershock sequence recorded there with more than 900 $M_L \geq 2$ aftershocks. Figure 16 shows the location in map view of this aftershock sequence, and the temporal fit of equation (10) to the cumulative number of events in the sequence yields $t_a \approx 0.24 \pm 0.05$ years. The complete PDF of t_a in plotted in the lower right part of Figure 16. Using equation (11), this value of t_a corresponds to $T_a = 1.5 \pm 0.32$ years, in fairly good agreement with the 1 year value estimated from the response of the MHT to seasonal variations of stress.

We proposed an analytical expression for the critical time T_a in *Ader et al.* (2013):

$$T_a = 2\pi\kappa \frac{a}{F(a,b)} \frac{D_c}{V_{pl}}, \tag{12}$$

where the function $F(a,b)$ of the fault parameters a and b refers to different models estimating the critical nucleation size, and can be $F(a,b) = a - b$ (*Ruina*, 1983), $F(a,b) = b$ (*Dieterich*, 1992), or $F(a,b) = b$ when $a/b < 0.37$ and $F(a,b) = \pi/2 \times (b-a)^2/b$ when $a/b > 0.5$ (*Rubin and Ampuero*, 2005). The parameter κ is of the order of a few units and represents the distance over which the deformation is accumulated at the onset of nucleation compared to the critical nucleation size (*Ader et al.*, 2013).

Unfortunately, equation (12) cannot lead to uncorrelated estimates of fault constitutive parameters, as it involves unknown parameters a, b, D_c and the function $F(a,b)$. However, even though the exact values of parameters a and b are still unknown, an order of magnitude of D_c can already be drawn from equation (12) by noting that both κ and $a/F(a,b)$ are most likely of the order of a few units, so that the factor $2\pi\kappa a/F(a,b) \sim 10^2$. Taking from the previous analysis $T_a \sim 1$ year and using the geodetically inferred convergence rate across the Nepal Himalaya of $V_{pl} \approx 20$ mm/yr, one can approximately say that D_c should be of the order of 0.1 mm. This estimate remains loosely constrained: additional analyses of the response of a finite fault to a stress perturbation still have to be undertaken in order to also take advantage of the information potentially contained in the amplitude and the phase of the response to eventually refine our knowledge of the constitutive fault parameters on the MHT from the Nepalese seismicity.

It should also be noted that these estimates implicitly assume fairly homogeneous fault properties, i.e., that the different faults on which earthquakes are produced have similar properties. A temporary seismic network, deployed from July to December 1995 above the midcrustal cluster of seismicity, showed that the microseismicity is localized in a volume, which approximately 10 kilometers thick and mainly localized within the overriding plate, around the down-dip end of the locked part of the MHT (*Cattin and Avouac*, 2000). The majority of these events are therefore produced on secondary faults around the MHT rather than on the MHT per se, and the study of the response of the midcrustal seismicity to stress perturbations yields averaged fault properties on these secondary faults. Extending these faults properties to the MHT therefore requires the hypothesis that the properties of these secondary faults are comparable to the fault properties of the MHT. Besides, it is noteworthy that the study suggests that the correlation of the timing of seismic events to perturbing stresses is due to the response of the growth of the nucleation zone to stress perturbations, and is not related to the event size, which only depends on the pre-stresses on the fault. According to these results, all events produced by the MHT should display a similar response to stress perturbations on the fault, regardless of their magnitude.

As has been described in the present introduction, the following four chapters present the four publications I wrote during my PhD. Chapter 1 uses geodetic observations of the deformation in Nepal to address the issue of the coupling pattern on the MHT and the convergence rate across the Nepal Himalaya, and assess some of the implications for seismic hazard in the Himalaya. The content of chapter 1 has been published in the Journal of Geophysical Research (*Ader et al.*, 2012a). The following chapter contains a paper submitted to Earth and Planetary Science Letters in January 2013 still under review, which examines the existence and amplitudes of periodicities in the seismicity rate in the Nepal Himalaya by developing a spectrum based on the Schuster test (*Ader and Avouac*, 2013). Chapter 3 focuses on the response of a rate-strengthening fault to harmonic shear stress perturbations through a publication in Geophysical Research Letters (*Ader et al.*, 2012b), which uses simple spring and slider systems with rate-strengthening rheology. Finally, chapter 4 presents the simulations on a finite 2D fault containing a rate-weakening patch using the BICYCLE algorithm to study the response of a rate-weakening fault to shear stress perturbations, together with results and conclusions.

Chapter 1

Convergence rate across the Nepal Himalaya and interseismic coupling on the Main Himalayan Thrust: Implications for seismic hazard

Published in Journal of Geophysical Research, VOL. 117, B04403, doi:10.1029/2011JB009071, 2012

Thomas Ader,[1,2] Jean-Philippe Avouac,[1] Jing Liu-Zeng,[3] Hélène Lyon-Caen,[2] Laurent Bollinger,[4] John Galetzka,[1] Jeff Genrich,[1] Marion Thomas,[1] Kristel Chanard,[1] Soma Nath Sapkota,[5] Sudhir Rajaure,[5] Prithvi Shrestha,[5] Lin Ding,[3] and Mireille Flouzat[4]

[1]Department of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA 91125, USA.

[2]Laboratoire de Géologie, Ecole Normale Supérieure, CNRS, 24 rue Lhomond, 75004 Paris, France.

[3]Key Laboratory of Continental Collision and Tibetan Plateau Uplift, Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing, People Republic of China.

[4]Commissariat à l'énergie Atomique, DAM, DIF 91297 Arpajon Cedex, France.

[5]National Seismological Centre, Department of Mines and Geology, Lainchaur, Kathmandu, Nepal.

Abstract

We document geodetic strain across the Nepal Himalaya using GPS times series from 30 stations in Nepal and southern Tibet, in addition to previously published campaign GPS points and leveling data and determine the pattern of interseismic coupling on the Main Himalayan Thrust fault (MHT). The noise on the daily GPS positions is modeled as a combination of white and colored noise, in order to infer secular velocities at the stations with consistent uncertainties. We then locate the pole of rotation of the Indian plate in the ITRF 2005 reference frame at longitude $= -1.34° \pm 3.31°$, latitude $= 51.4° \pm 0.3°$ with an angular velocity of $\Omega = 0.5029 \pm 0.0072°/\text{Myr}$. The pattern of coupling on the MHT is computed on a fault dipping 10° to the north and whose strike approximately follows the arcuate shape of the Himalaya. The model indicates that the MHT is locked from the surface to a distance of approximately 100 km down dip, corresponding to a depth of 15 to 20 km. In map view, the transition zone between the locked portion of the MHT and the portion which is creeping at the long term slip rate seems to be at the most a few tens of kilometers wide and coincides with the belt of midcrustal microseismicity underneath the Himalaya. According to a previous study based on thermokinematic modeling of thermochronological and thermobarometric data, this transition seems to happen in a zone where the temperature reaches 350°C. The convergence between India and South Tibet proceeds at a rate of 17.8 ± 0.5 mm/yr in central and eastern Nepal and 20.5 ± 1 mm/yr in western Nepal. The moment deficit due to locking of the MHT in the interseismic period accrues at a rate of $6.6 \pm 0.4 \times 10^{19}$ Nm/yr on the MHT underneath Nepal. For comparison, the moment released by the seismicity over the past 500 years, including 14 $M_W \geq 7$ earthquakes with moment magnitudes up to 8.5, amounts to only 0.9×10^{19} Nm/yr, indicating a large deficit of seismic slip over that period or very infrequent large slow slip events. No large slow slip event has been observed however over the 20 years covered by geodetic measurements in the Nepal Himalaya. We discuss the magnitude and return period of M>8 earthquakes required to balance the long term slip budget on the MHT.

1.1 Introduction

Most of the convergence rate across the Himalaya of central Nepal has been shown to be absorbed by slip along a major basal thrust fault, the Main Himalayan Thrust fault (MHT), which emerges at the surface along the front of the Himalayan foothills (e.g., review by *Avouac*, 2003). It is therefore a good first approximation to assume that the Himalaya overthrusts the Indian crust with little internal deformation. A corollary is that large earthquakes which are known to recur along the Himalayan front (*Ambraseys and Douglas*, 2004; *Bilham*, 2004) must be associated with ruptures of the MHT. The largest Himalayan earthquake which has occurred in the instrumental period in the Himalaya is the 1950 Assam earthquake, which reached a moment magnitude estimated to $M_W \sim 8.5$ (*Ambraseys and Douglas*, 2004; *Chen and Molnar*, 1977). Paleoseismological investigations suggest that even larger earthquakes may have happened in the past (*Lavé et al.*, 2005; *Kumar et al.*, 2006, 2010). Some studies have also suggested some out of sequence thrusting, with possible reactivation of thrust faults in the Main Central Thrust (MCT) zone (e.g., *Hodges et al.*, 2004; *Seeber and Gornitz*, 1983).

In this context the pattern of geodetic strain measured across the Nepal Himalaya over the last 20 years, during which no large earthquake has happened (the largest earthquake during this period has a moment magnitude of 5.6), may be used to determine the pattern of locking of the fault in the interseismic period and estimate the return period of large earthquakes required to release the elastic strain which builds up in the interseismic period. This pattern is quantitatively characterized by the 'interseismic coupling ratio', defined as the ratio of the deficit of slip rate in the interseismic period divided by the long term slip rate.

The approach used here is commonly used to study subduction zones. It has been shown that in the subduction context the pattern of locking is generally very heterogeneous and shows a correlation with the rupture areas of large interplate eartquakes (*Chlieh et al.*, 2008; *Moreno et al.*, 2010; *Suwa et al.*, 2006; *Ozawa et al.*, 2011; *Loveless and Meade*, 2010; *Freymueller et al.*, 2000; *Wallace et al.*, 2004).

Modeling of interseismic strain may reveal creeping patches that could act as barriers to the propagation of large earthquakes, as well as locked asperities, which might hence help determine the possible extent of future seismic ruptures (*Bürgmann et al.*, 2005; *Kaneko et al.*, 2010). Hereafter, we introduce the seismotectonic setting of the Nepal Himalaya in section 1.2 and we describe in section 1.3 the processing applied to the geodetic data used in this study, as well as the computation of the pole of rotation of the Indian Plate in the ITRF 2005 reference frame. The supplementary material details the processing from raw GPS data to estimates of secular velocities at each station with coherent uncertainties. The coupling pattern on the MHT inferred from geodetic data is presented in section 1.4, and then used in section 1.5 to evaluate the seismic hazard in Nepal. The supplementary material contains details on the inversion of the geodetic data, such as resolution and smoothing method, as well as a discussion on the modeling of the extension of the Tibetan plateau. It also derives a couple additional estimates of the recurrence time of large earthquakes.

1.2 Seismotectonic Setting

Most of the crustal deformation in the Himalaya occurs on the Main Himalayan Thrust fault (MHT) (e.g., *Cattin and Avouac*, 2000; *Lavé and Avouac*, 2000), where the Indian lithosphere underthrusts beneath the chain (*Zhao et al.*, 1993). The MHT absorbs about 20 mm/yr of the India-Eurasia convergence (*Bilham et al.*, 1997), which accounts for about half of the total convergence rate between the Indian and Eurasian plate (*Bettinelli et al.*, 2006). The MHT reaches the surface at the Main Frontal Thrust fault (MFT) (*Nakata*, 1989), where the secular slip rate has been estimated from the study of uplift of Holocene terraces to be 21.5 ± 1.5 mm/yr in central Nepal (*Lavé and Avouac*, 2000) and 19 ± 6 mm/yr in western Nepal (*Mugnier et al.*, 2003). Previous geodetic studies (*Bilham et al.*, 1997; *Jouanne et al.*, 1999, 2004; *Larson et al.*, 1999; *Bettinelli et al.*, 2006) indicate that the MHT is actually locked at the surface and roots about 100 km to the north of the MFT into a subhorizontal shear zone of probably thermally enhanced ductile flow (*Cattin and Avouac*, 2000). A fraction of geodetic interseismic strain could be due to anelastic deformation (*Bilham et al.*, 1997; *Meade*, 2010). This fraction is probably small given that the slip rate on the MFT matches the shortening rate across the range, and is therefore neglected in this study. The locked portion of the fault elastically absorbs the 20 mm/yr of shortening across the Nepal Himalaya during the interseismic period, and releases this deformation during large ($M_W > 8$) earthquakes (*Molnar*, 1987; *Bilham et al.*, 1995; *Avouac et al.*, 2001). The observation of meter-scale displacements on some regions of the MFT indicates that during those large earthquakes, the locked portion of the fault sometimes ruptures all the way to the surface (*Nakata*, 1989; *Lavé et al.*, 2005). This description of the seismic cycle in the Himalaya is supported by the observation of a belt of microseismicity at the creeping-locked transition (*Pandey et al.*, 1995, 1999), which underlies a zone of greater stress accumulation

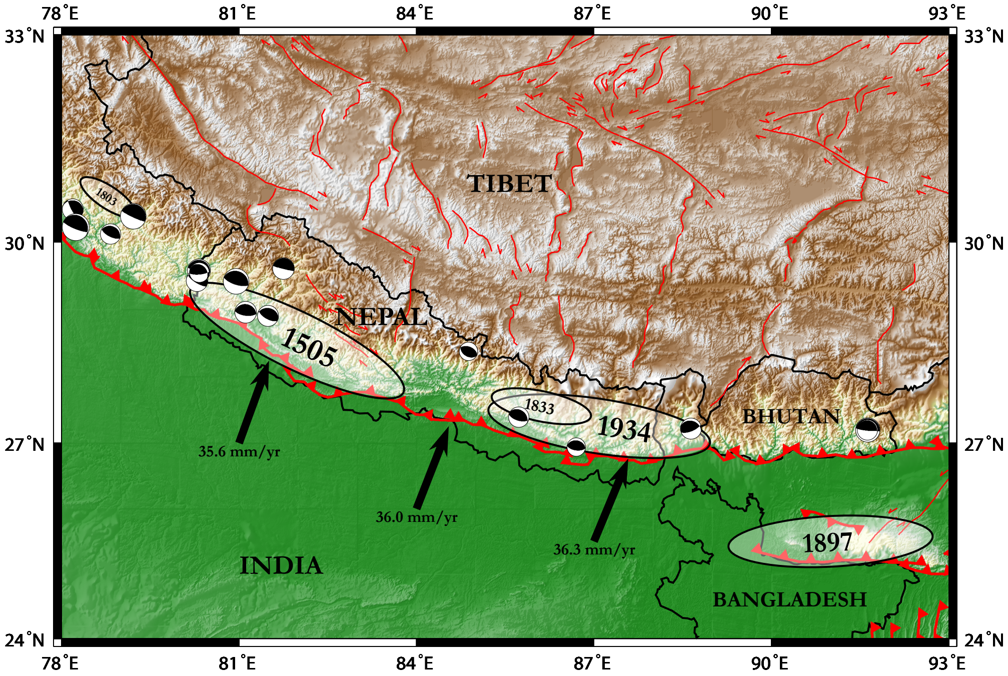


Figure 1.1: Seismotectonic setting of the Himalaya. Arrows show Indian plate motion relative to Eurasia computed using the rotation poles of Eurasian plate in ITRF 2005 from *Altamimi* (2009), and Indian plate in ITRF 2005 from this study. Focal mechanisms show thrust events (rake = 90° ± 45°) from the CMT catalog between 1976 and 2011. White ellipses show locations of historical earthquakes according to *Ambraseys and Douglas* (2004). Ellipses sizes are scaled with the earthquakes magnitudes, and might not represent reliably the area ruptured during these earthquakes. Active faults (in red) map modified from *Styron et al.* (2011).

rate (*Cattin and Avouac*, 2000; *Bollinger et al.*, 2004). An accurate knowledge of the coupling pattern of the MHT is thus a paramount element to investigate the distribution of the seismicity in time, space and magnitude, relating it to the accumulation of moment deficit on the MHT.

Previous geodetic studies using GPS campaign and leveling measurements (*Jouanne et al.*, 2004; *Bollinger et al.*, 2004) as well as data from a few continuous GPS stations (*Bettinelli et al.*, 2006) assumed a fault locked from the surface to a certain depth and found a satisfying fit to the data with a fault dipping about 10° to the north and a downdip end of the locked part of the fault about 100 km along dip from its surface trace. The dataset was insufficient to resolve details of the interseismic coupling pattern, including possible along-strike variations. Here we take adavantage of an improved geodetic dataset which we combine with previous geodetic measurements to better resolve the pattern of coupling on the MHT and investigate the relationship between geodetic strain and seismicity in the Himalaya.

1.3 Data used to determine the coupling pattern on the MHT and the convergence rate

1.3.1 Continuous GPS stations

We use data from a network of continuous GPS stations currently consisting of 24 stations, spanning the Nepalese territory (supplementary Figure S1). In addition, a station in Sarangkot (SRGK) has been in operation from March 2005 to February 2007, providing one additional velocity point (see supplementary table S2 for velocities and coordinates of the stations as well as their dates of operation), bringing up the total number of continuous GPS velocities available in Nepal to 25. The station SIMR in Simara was in operation from November 1997 until April 2005, but was eventually replaced by the nearby station SIM4. The stations BRNG and MSTG respectively in Biratnagar and Lo-Mantang (Upper Mustang) are now destroyed and have been replaced by the 2 stations BRN2 and MST2 in 2009. We also used data from 5 continuous GPS stations in southern Tibet, provided by the Institute of Tibetan Plateau Research, Chinese Academy of Science.

Raw dual frequency code and phase observations are translated to RINEX files. Observations were taken at 30 second intervals, or downsampled to that rate if the sampling rate was higher (15 seconds) and then processed with the GAMIT/GLOBK software package (*Herring et al.*, 2009). Daily network solutions include neighboring IGS sites (Table S1). The daily regional solutions are combined with solutions for five global IGS networks (SOPAC) to yield daily station coordinates for all sites in the ITRF2005 realization (*Altamimi*, 2009).

Previous studies have shown that the dominant sources of signal in the GPS time series are the interseismic secular rate of loading and seasonal variations due to surface load variations induced by the Monsoon regime (*Bettinelli et al.*, 2006, 2008). However a number of artificial steps can also appear in the time series, mostly due to maintenance operations, small local earthquakes or equipment malfunction. The GPS position time series are therefore modeled as follows:

$$x(t) = p_1 + p_2 t + \sum_{i=1}^{4} \left(p_{2i+1} \cos 2\pi \frac{t}{T/i} + p_{2i+2} \sin 2\pi \frac{t}{T/i} \right) + \sum_{i=1}^{S} p_{10+i} \mathbf{1}_{t>t_i^s}, \qquad (1.1)$$

where $T = 1$ year, S is the total number of steps known to be in the time series, t_i^s is the time at which each step occurs and the linear parameters p_i are defined as

- a constant offset (parameter p_1),
- the secular velocity (parameter p_2),
- annual variations, modeled by sine waves of periods of 1, 1/2, 1/3 and 1/4 years (parameters p_3 to p_{10}),
- steps in the time series (for s steps, parameters p_{11} to p_{10+s}).

The function $\mathbf{1}_{t>t_i^s}$ is defined as follows:

$$\mathbf{1}_{t>\tau_i} = \begin{cases} 0 & \text{if } t < t_i^s \\ 1 & \text{if } t \geq t_i^s \end{cases} . \qquad (1.2)$$

The linear parameters p_i are estimated through a standard least-squares inversion.

The nominal uncertainties on the daily positions do not account for all sort of additional sources of signal not taken into account in this decomposition (steps too small to be detected for example). Also the daily estimates are temporaly correlated and it is therefore incorrect to assume a purely white noise model (*Zhang et al.*, 1997; *Langbein and Johnson*, 1997; *Williams*, 2003a). For these reasons we have determined a noise model and the related covariance matrix following the approach of *Williams* (2003a) and *Williams et al.* (2004). Details are given in Supplements.

1.3.2 GPS campaign measurements

Several GPS campaigns have been conducted, starting in 1991 with the CIRES network, that covers Nepal from the Higher Himalaya to the Himalayan foreland (*Bilham et al.*, 1997). Then, starting in 1995, the LDG campaign focused on points at the longitude of Katmandu while the IDYLHIM program, including some of the CIRES points and adding new ones, was designed to study the Himalaya of central and western Nepal (*Jouanne et al.*, 2004). We also used recently published GPS campaign measurement from the Garhwal-Kumaon Himalaya (India) close to the far western border of Nepal (*Ponraj et al.*, 2011).

1.3.3 Determination of the Euler pole of the Indian plate in the ITRF2005 reference frame

In order to invert the geodetic data for the pattern of coupling on the MHT, we first need to express the GPS velocities with respect to the Indian plate reference frame. In this section, we compute the coordinates and angular velocity of the Indian plate's Euler pole in the ITRF 2005 reference frame that we will use to put our data in the Indian reference frame. We use the secular velocities computed following the method described in the previous section of 4 stations from southern Nepal (DNGD, NPGJ, SIMR/SIM4, BRNG), the IGS station IISC, the two stations HYDE and MALD as well as the DORIS station COLA in Columbo. We also use the velocities at 12 Indian stations published by *Banerjee et al.* (2008). The time series at those stations being not available we could not estimate the uncertainties at those stations using the approach described in this study, used at the other continuous sites. The uncertainties on the velocities obtained with the Maximum Likelihood Estimation (MLE) algorithm used in this study (see supplements for details) might differ from the uncertainties provided by *Banerjee et al.* (2008). However, we observe that at the 3 common stations HYDE, MALD and IISC, the velocities match within their uncertainties, which are 1 to 1.5 times larger in the estimates of *Banerjee et al.* (2008). The Indian plate is regarded as a rigid plate, in first order agreement with the conclusion of *Banerjee et al.* (2008) who inferred a 2 ± 1mm/yr north-south shortening across the Indian subcontinent (i.e., strain rate less than $10^{-14}s^{-1}$). The observed GPS velocities are thus entirely modeled by the rotation of a rigid plate. Adding data from southern Nepal insures us to be in the northern India reference frame in case of a slight north-south shortening of the Indian plate. This best fit model is plotted on Figure 1.2 and corresponds to the following Euler pole describing the rigid Indian plate motion in the ITRF05 reference frame:

$$\begin{cases} \text{longitude} = -1.34^\circ \pm 3.31^\circ \\ \text{latitude} = 51.4^\circ \pm 0.3^\circ \\ \Omega = 0.5029 \pm 0.0072^\circ/\text{Myr} \end{cases} . \tag{1.3}$$

Those parameters are in good agreement with previously published GPS-based models of the Indian plates motion (*Socquet*, 2003; *Bettinelli et al.*, 2006; *Banerjee et al.*, 2008), as is shown in Figure 1.2(c).

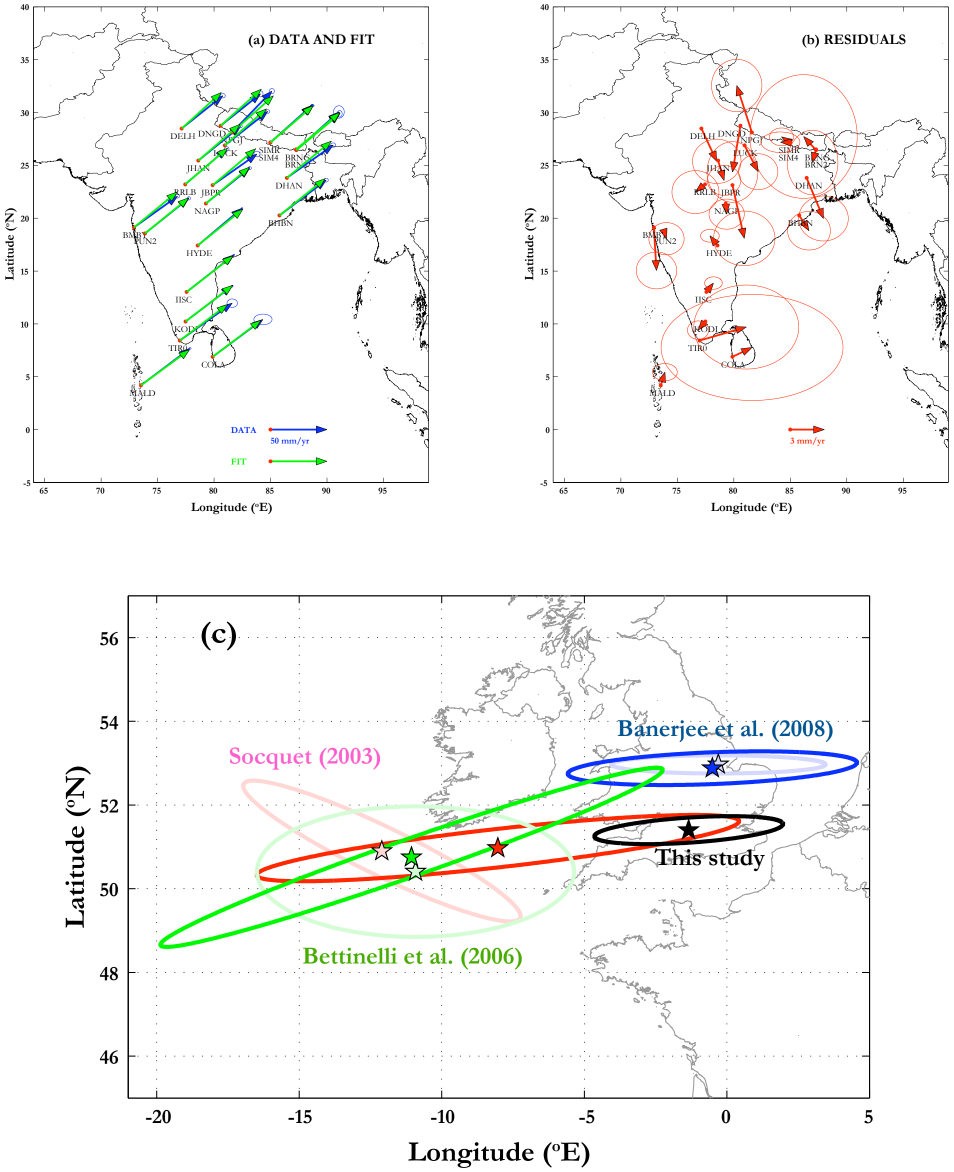


Figure 1.2: Pole of rotation of the rigid Indian plate in the ITRF 2005 reference frame. (a) Fit to the GPS velocities assuming a rigid Indian plate, (b) residuals of the fit, (c) Positions of Euler poles describing the rotation of the Indian Plate in the ITRF 2005 reference frame from the literature (light colors) and recomputed in this study with the original published data (corresponding dark colors). The studies from which each Euler pole is taken are indicated in the figure.

1.3.4 Leveling data

The most reliable vertical velocities available are the leveling data collected between 1977 and 1990 by the Survey of Nepal (DMG) along the Birganj-Katmandu-Kodari road, in central Nepal (line of white dots on Figure 1.3(a)) (*Jackson and Bilham*, 1994). The GPS vertical velocities have large uncertainties (table S2). They are used in the inversion presented below although we find that they do not add much constraint to the model (Figure 1.3(b)).

1.4 Coupling method and results

1.4.1 Inversion method

We invert the geodetic data using the backslip modeling approach (*Savage*, 1983): the long term slip along the fault (related to the long term convergence between India and southern Tibet as it is assumed that the hanging wall and footwall do not deform in the long term) is subtracted from the interseismic geodetic displacements, and we thus solve for a backward slip which represents interseismic locking of the MHT. Displacements at the surface are related to fault slip at depth assuming a linear purely elastic half space (Okada, 1985). The backslip formulation is rigorously correct only in the case of a purely planar fault (*Vergne et al.*, 2001). Our fault model dips straight 10° towards the north. It should be noticed that for a locked fault the geometry of the modeled fault does not need to reproduce the geometry of the real fault provided they match at the downdip end of the locked zone (*Vergne et al.*, 2001). Especially, the ramp on the MHT that has been reported in numerous studies (e.g., *Pandey et al.*, 1995) does not need to be modeled here as it falls within the locked zone of the MHT, as will be seen later. In order to account for the arcuate shape of the Nepal Himalaya we consider 3 segments, whose directions schematically follow the front of the Himalaya (figures 1.3(a), 1.5(a) and 1.7). Along the MFT, the first segment covers longitudes from 78.4°E to 82.4°E, the second segment goes from 82.4°E to 86.6°E and the third one from 86.6°E to 88.1°E. The fault is then discretized into 935 rectangular patches (17 along dip, 55 along strike) of about 20×15 km, for each of which the slip is computed. This back-slip velocity is then used to estimate the interseismic coupling (ISC) which quantifies the degree of locking of the fault:

$$\mathrm{ISC} = \frac{\text{deficit of slip rate on the patch}}{\text{long term slip rate}}. \tag{1.4}$$

In principle, in absence of transient slip events along the MHT, interseismic coupling should be between 0 and 1. A coupling of 0 indicates that the patch creeps at the long term slip rate, and a coupling of 1 indicates that the patch is locked. Negative values of the coupling would imply that the fault could creep faster than the far field velocity, while coupling values greater than 1 would mean that the patch is creeping backward. The coupling is thus constrained to be between 0 and 1 in this inversion.

The Tibetan plateau is also undergoing some east-west extension through a network of north-south oriented grabens and east-west strike-slip faults (*Armijo et al.*, 1986) with a detectable geodetic signal (*Chen et al.*, 2004; *Styron et al.*, 2011). Here we model this extension by letting the far field velocity vary both in norm and azimuth in eastern and western Nepal. The separation between the east and the west is chosen so that it coincides with the Thakkola graben in the Himalaya (dashed line on Figure 1.3(a)), which is the only graben that extends into the Himalaya. On top of the east-west separation, we model the cumulative contribution of the other grabens by a linear extension along the N98.2E direction applied to all stations north of the northern border of Nepal (those being the only stations displaying an obvious deviation), justified by the observation by *Styron et al.* (2011) that the arc-parallel extension is uniformly distributed throughout the Nepal Himalaya. This direction has been selected because it minimizes the projection of the velocity correction due to the extension onto the far field velocity in southern Tibet. Adding both effects (the graben and the linear extension in southern Tibet) reduced the variance of the fit by an amount that has a probability of 80% to be significant, according to the F-test (*Press et al.*, 1992). As will be specified later, the model used for the arc-parallel extension actually has little impact on the quantities estimated in this study.

Assuming an elastic medium, the surface displacements at the stations are related to the slip distribution on the fault through the linear equation:

$$\mathbf{d} = \mathbf{G}\mathbf{m}, \tag{1.5}$$

where $\mathbf{d}$ is the data vector made of the surface displacements at the measurement points, $\mathbf{G}$ is the Green's functions matrix computed using the semi-analytical formulation published by *Okada* (1985), and $\mathbf{m}$ is the vector of parameters we are looking for (strike and dip slip on the 935 15×20 km rectangular sub-patches used to mesh the fault, the long term convergence rates across western and eastern Nepal Himalaya and the linear extension rate of southern Tibet). The displacements on the fault are determined from a standard least-squares inversion.

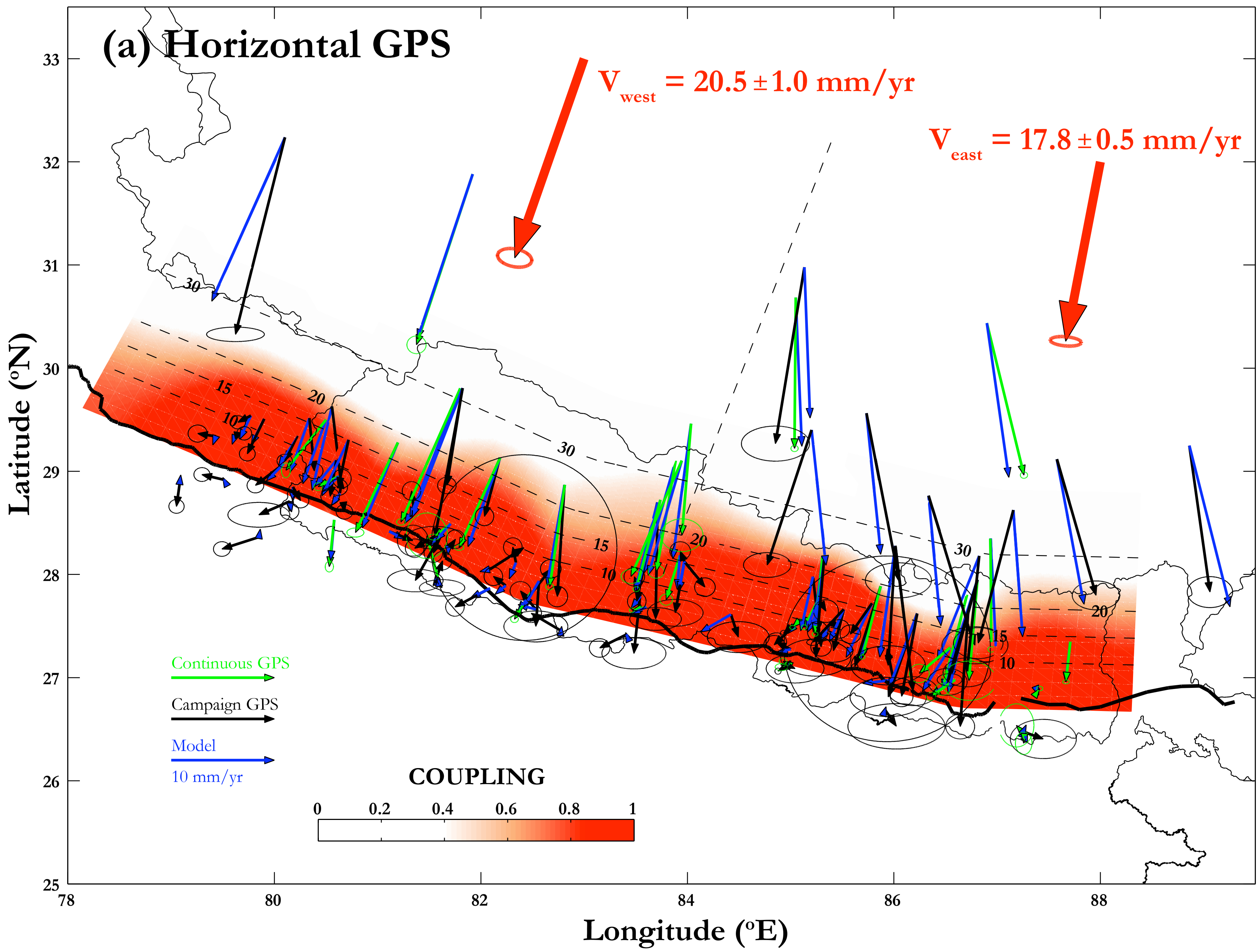
Owing to the sparsity of geodetic data, the problem is underdetermined. In order to regularize the problem, we apply a Laplacian smoothing to the slip on the fault, weighting it according to the resolution on each patch (see supplementary section S.3 for more details). This adds two Laplace equations per patch (one for slip in the dip direction, and one for slip in the strike direction), which, together with the geodetic data to fit, make the problem overdetermined. The east-west extension of the Tibetan plateau causes the velocities on the plateau to vary from east to west but also to deviate from being purely dip-slip. If no additional constraint is imposed, this divergence is compensated by a strike-slip component at the downdip end of the fault, which is an artifact that we are not interested in modeling in this study. To counter this effect, we require the solution to minimize the coupling at downdip end of the fault.

Summarizing those constraints into an equation, the slip on the fault (vector $\mathbf{m}$) is obtained by minimizing the quantity:

$$\chi^2(\mathbf{m}) = \parallel \mathbf{C_d}^{-1/2}(\mathbf{Gm} - \mathbf{d}) \parallel^2 + \lambda \parallel \mathbf{\Lambda m} \parallel^2 + \mu \parallel \mathbf{I_d m} \parallel^2, \tag{1.6}$$

where $\mathbf{C_d}$ is the data covariance matrix, $\mathbf{\Lambda}$ is the Laplacian matrix and $\mathbf{I_d}$ is the matrix which, when multiplied by $\mathbf{m}$ returns the components of $\mathbf{m}$ corresponding to the slip at depth, λ and μ measure the weights attributed to each constraint. The parameter μ is manually adjusted ($\mu = 0.01$) to prevent any significant strike-slip component at the downdip end of the fault without affecting the slip on the rest of the fault, while the parameter λ is chosen to minimise χ^2, which leads to $\lambda = 0.85$ (Figure 1.8(a)). The values of the slip $\mathbf{m}$ are then divided by the corresponding value of the long term slip rate to obtain interseismic coupling.

The uncertainties have been multiplied by 5 at stations that displayed an abnormal behavior, that would require additional "parameters" not included in our model to be explained. This is the case for instance of NPGJ that has a northward motion of about 3.5 mm/yr, whereas it sits about 20 km south of the MFT, and should hence have almost no northward motion.


(a) Horizontal GPS
V_{west} = 20.5 ± 1.0 mm/yr
V_{east} = 17.8 ± 0.5 mm/yr
30
20
15
10
Continuous GPS
Campaign GPS
Model
10 mm/yr
COUPLING
0
0.2
0.4
0.6
0.8
1
Latitude (°N)
Longitude (°E)
25
26
27
28
29
30
31
32
33
78
80
82
84
86
88

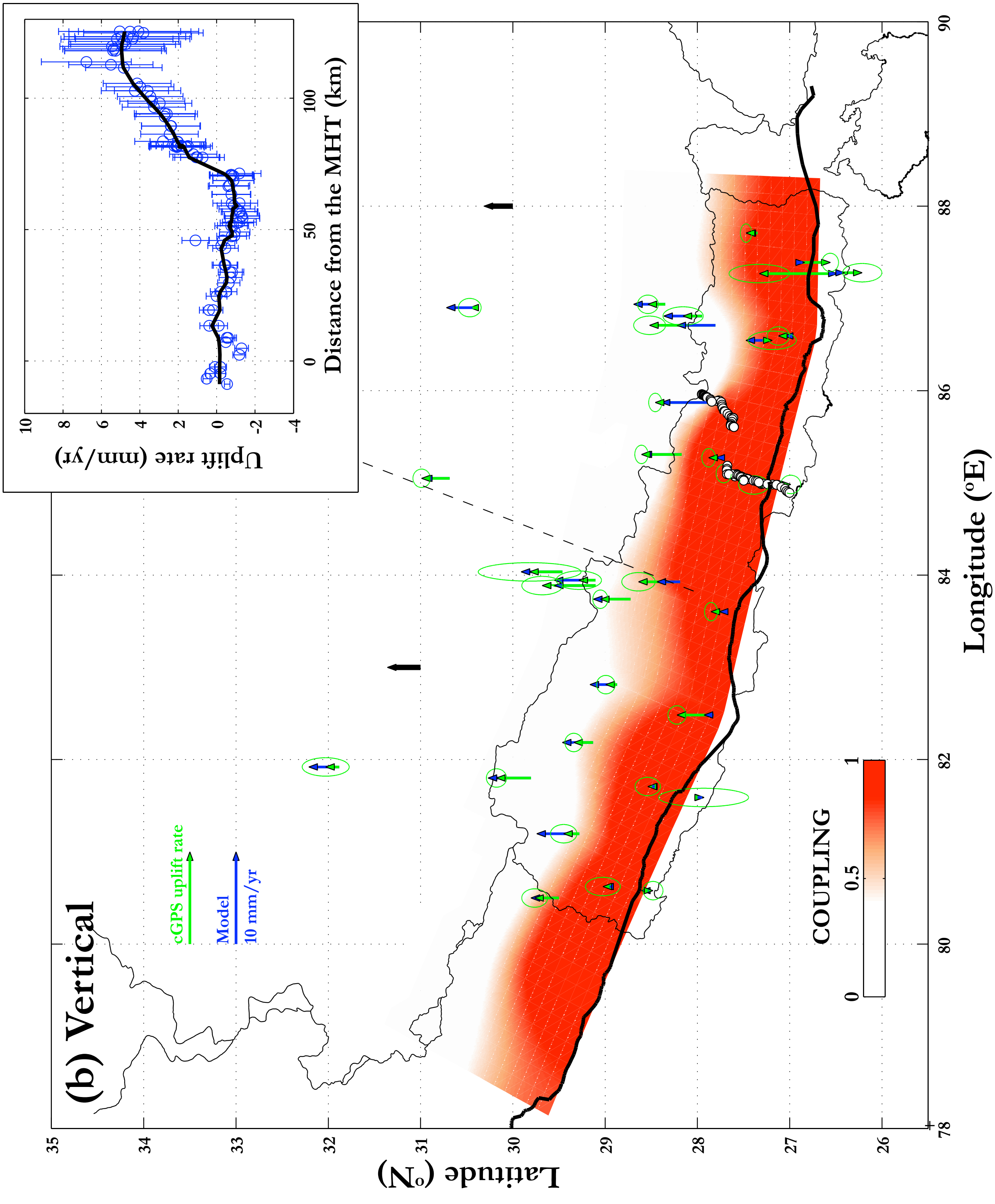

(b) Vertical
cGPS uplift rate
Model
10 mm/yr
COUPLING
0
0.5
1
Latitude (°N)
Longitude (°E)
35
34
33
32
31
30
29
28
27
26
78
80
82
84
86
88
90
Uplift rate (mm/yr)
Distance from the MHT (km)
10
8
6
4
2
0
-2
-4
0
50
100

Figure 1.3: Fit to the geodetic data. (a) Comparison between observed and predicted horizontal velocities. Interseismic coupling is shown as shades of red on the fault. The GPS data with corresponding error bars are plotted respectively as green and black arrows for the continuous and campaign GPS measurements. Blue arrows show predicted velocities according to the plotted pattern of interseismic coupling. Dashed line approximately trending north-south indicates the east-west separation, on each side of which the secular velocity can be different. Red arrows represent the east and west long term convergence rate across the Himalaya. Black dashed lines with numbers represent contour lines of fault depth (in km). (b) Fit to the continuous GPS vertical velocities (map) and to the spirit leveling data (inset). White dots show location of the leveling line. The inset shows the fit to leveling data.

1.4.2 Results of the inversion

The pattern of interseismic coupling and the convergence rates across the Himalaya, east and west of the Thakkola graben, determined from the inversion of the GPS velocities and leveling line are plotted in Figure 1.3. The reduced chi squares of the fit to the different data sets are given in Table 1.1. These values show that there is no need to renormalize the uncertainties as all reduced chi-squares are of the order of unity.

Data set	reduced χ^2
continuous GPS	1.92
campaign GPS	2.94
leveling	1.69
total	2.29

Table 1.1: Values of the reduced χ^2 of the fit to the different data sets.

The overall reduced chi-square value obtained for this best fitting model is 2.29 suggesting that some small fraction of the signal might not be adequately explained by the model. However, the residuals plotted on Figure 1.4 show no systematic misfits, indicating that no significant signal has been left out by the model. The somewhat large residuals at the north-westernmost station might be due to the effect of the Karakoram fault.

The resolution at each of the fault's patches is plotted in Figure 1.5(a) (see supplementary material section S.3 for details on how resolution is determined). The resolution is expressed here in terms of the characteristic size of smallest inhomegeneities of coupling which could in principle be resolved given the spatial distribution and the uncertainties of the measurements. For clarity, we saturated Figure 1.5(a) at a resolution of 80km, since one can assume that above such a value there is simply no resolution on the corresponding

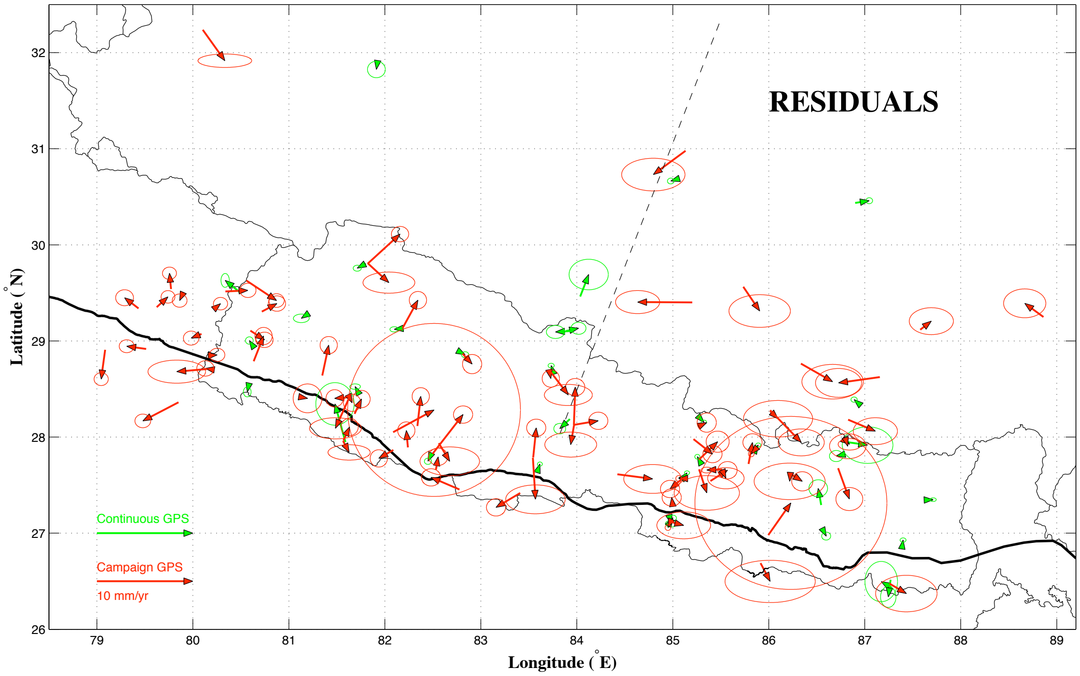


Figure 1.4: Residuals of the fit to the geodetic data. The GPS residuals with corresponding error bars are plotted as red arrows for the campaign measurements and as green arrows for continuous GPS. The dashed line represents the position of the Thakkola graben.

patch and that the slip on this patch is entirely determined by the slip on the neighboring patches. Those patches with no resolution are shaded in grey in Figure 1.5. The resolution on the fault can be as good as 20 km close to the MFT (i.e., at shallow depth) and gradually increases to about 60 km at greater depth along the MHT. On the edges, there is no resolution, indicating that there would be no point extending the fault along strike.

A striking result of this inversion is that the fault seems to be fully locked from where it emerges at the surface along the Himalayan foothills to beneath the front of the high range about 100 km to the north. At the resolution afforded by this inversion, no zone of creep appears close to the surface. Especially, interseismic coupling appears to be very homogeneous along strike. The subtle along strike inhomogeneities of the coupling pattern probably mostly reflect inhomogeneities of the resolution. It is noteworthy that the width of the transition zone at the downdip end of the locked fault zone is typically of the order of the nominal resolution size. It is sharpest in the area where the resolution is enhanced by the leveling data. There, the transition from a fully locked fault to a fault creeping at the long term slip rate occurs within 20 km according to our coupling model (red line on Figure 1.6) but it could in reality be even sharper. Elsewhere the transition is always wider (light red shaded curve on Figure 1.6), probably because of the more limited resolution of

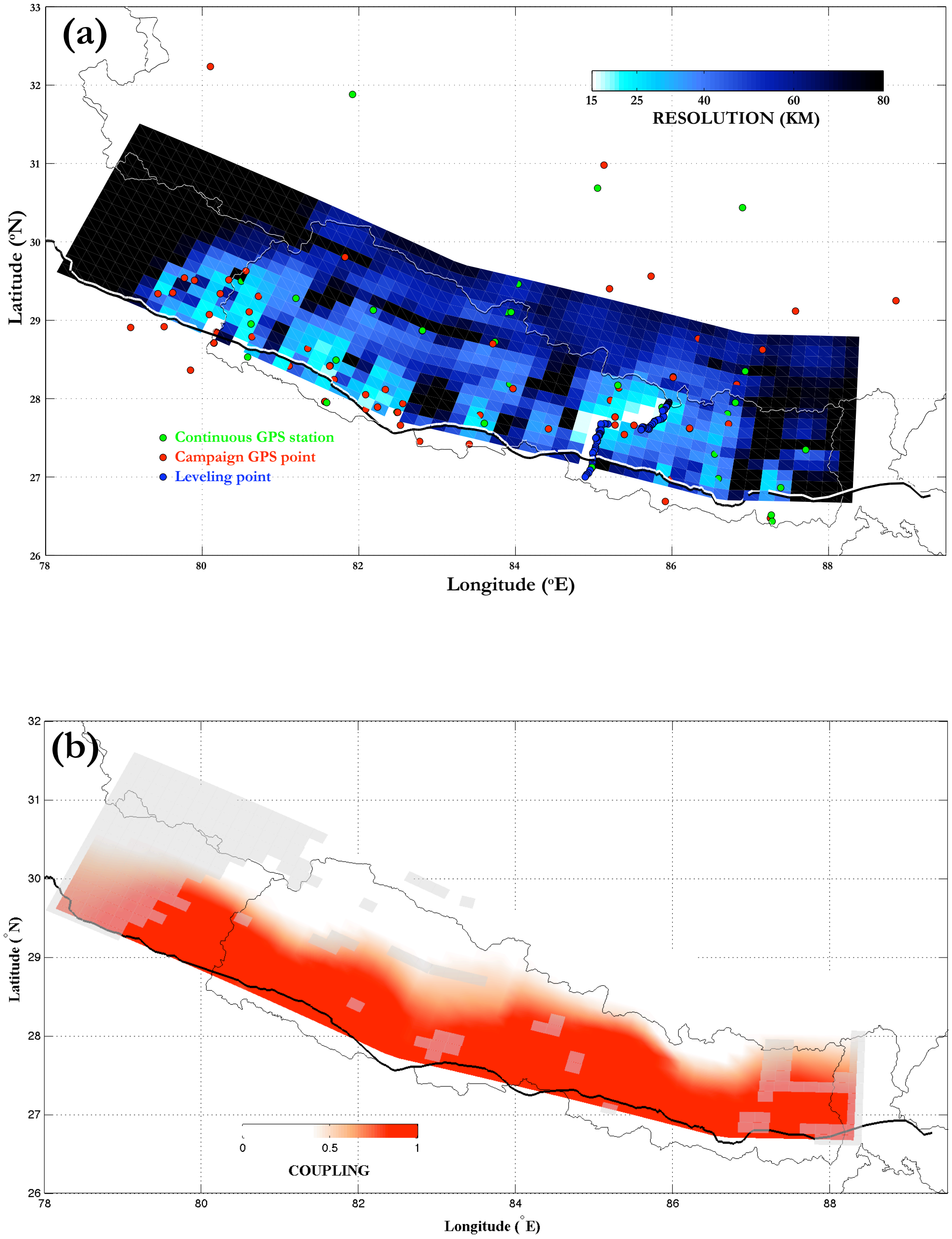


Figure 1.5: (a) Resolution on each of the patches of the fault, given in km. See text and supplements for details on the computation of the resolution. Locations of the data points used to compute the resolution are indicated in the figure and legend. (b) Map of the coupling on the MHT on which patches with no resolution (resolution > 80 km) are masked.

the inversion. A thorough reader will notice slight decreases of coupling at the junctions between the planes used to model the fault. Those are a pure artifact of the model, and moving the location of those junctions also moves the slight decrease of coupling.

The convergence rate across the Himalaya is estimated to 17.8 ± 0.5 mm/yr east of the Thakkola graben and 20.5 ± 1 mm/yr west of it and the annual moment deficit to $6.6 \pm 0.4 \times 10^{19}$ Nm/yr, assuming a shear modulus of 30 GPa. Those uncertainties are given at the 1-σ confidence level, as is the case for all uncertainties reported throughout this paper. One should also keep in mind that they are computed using the 1-σ uncertainties on geodetic data but that they do not include error on the model itself, other than being rescaled in order to have a final χ^2 of the fit equal to 1. As a result, those uncertainties are slightly underestimated.

The extension rate across the Thakkola graben is estimated to 3.1 ± 2.6 mm/yr. Elsewhere, we get a distributed extension rate of the southern Tibetan plateau of 9.3 ± 2.1 mm.yr^{-1}.10^{-3}km^{-1} (or nstrain.yr^{-1}). This adds up to a total extension rate of the southern Tibetan plateau north of Nepal of 12.4 ± 4.7 nstrain.yr^{-1}, consistent with, although somewhat lower than the ~ 35 nstrain.yr^{-1} estimated by *Styron et al.* (2011). This extension is accommodated by north-south grabens and east-west strike-slip faults in southern Tibet. One must keep in mind that those rates highly depend on the assumed direction of extension, which was chosen in order to interfere as little as possible with our results of coupling on the MHT and consequently these rates should be regarded with caution. The coupling pattern on the MHT and the important quantities estimated in this study (convergence velocity and moment deficit accumulation rate) are not much affected by the model used for the extension of the Tibetan plateau. Not adding the linear extension and letting the strike slip component on the MHT being as large as needed results in long term velocities of 18.1 ± 0.5 mm/yr east of the Thakkola graben and 20.8 ± 1 mm/yr west of it, and in an annual moment deficit of $6.7 \pm 0.4 \times 10^{19}$ Nm/yr. Those values are less than 2% different from the ones derived with the model of linear arc parallel extension in southern Tibet. Figure S5 shows that the direction selected for the arc-parallel extension does not have an significant impact either.

1.5 Implications

1.5.1 Convergence rate across the Himalaya

The estimates of the geodetic convergence rate across the Himalaya obtained in this study, 17.8 ± 0.5 mm/yr and 20.5 ± 1 mm/yr east and west of the Thakkola graben respectively, are better constrained than but consistent with previous estimates, e.g., 19 ± 2.5mm/yr in central and eastern Nepal according to *Bettinelli et al.* (2006) or 16 mm/yr in eastern Nepal according to *Banerjee et al.* (2008). Those rates are also close to the 21.5 ± 1.5 mm/yr of rate determined by *Lavé and Avouac* (2000) from deformed Holocene terraces. This observation indicates that decadal geodetic interseismic deformation of the upper crust is essentially elastic and entirely released over the longer term by localized slip along the MHT, without any significant shortening of the hanging wall. Especially, over the past decade, no deformation within the MCT zone is required to explain the data.

We also observe that the convergence is nearly perpendicular to the strike of the range front and parallel to the azimuth of slip vectors on thrust faults along the Himalayan arc as well as to the stretching lineation observed in the Lesser Himalaya (*Bollinger et al.*, 2004). This observation still holds if we do not resort to the linear arc-parallel extension of the Tibetan plateau and let the strike slip on the fault be as large as needed to fit the geodetic data. The direction of convergence across the range must have been remarkably stationary at the 10-15 Ma time scale of the development of the Lesser Himalayan duplex system (*Bollinger et al.*, 2004).

1.5.2 Temperature control on the downdip end of the Locked Fault Zone

Along-strike variations of geodetic strain across the Nepal Himalaya are thus small and are accounted for by relatively minor variations of the location of the downdip end of the locked fault zone (Figure 1.5). Figure 1.6 shows along-dip variations of interseismic coupling along the MHT at the location of the Kathmandu basin where the resolution is best, as well as the coupling calculated on the rest of the fault. The seismicity rate is also indicated on this figure as a bar plot and appears to peak in the zone where the coupling values drop. Given the cylindrical geometry of Himalayan structure, this observation makes it difficult to identify what factor primarily controls the downdip extent of the locked fault zone. Laboratory experiments on quartzo-feldspathic rocks show frictional sliding transitions from rate-weakening, favoring unstable slip sliding, to rate-strengthening, favoring stable creep, at a temperature around 350°C (*Blanpied et al.*, 1995; *Marone*, 1998). This has been

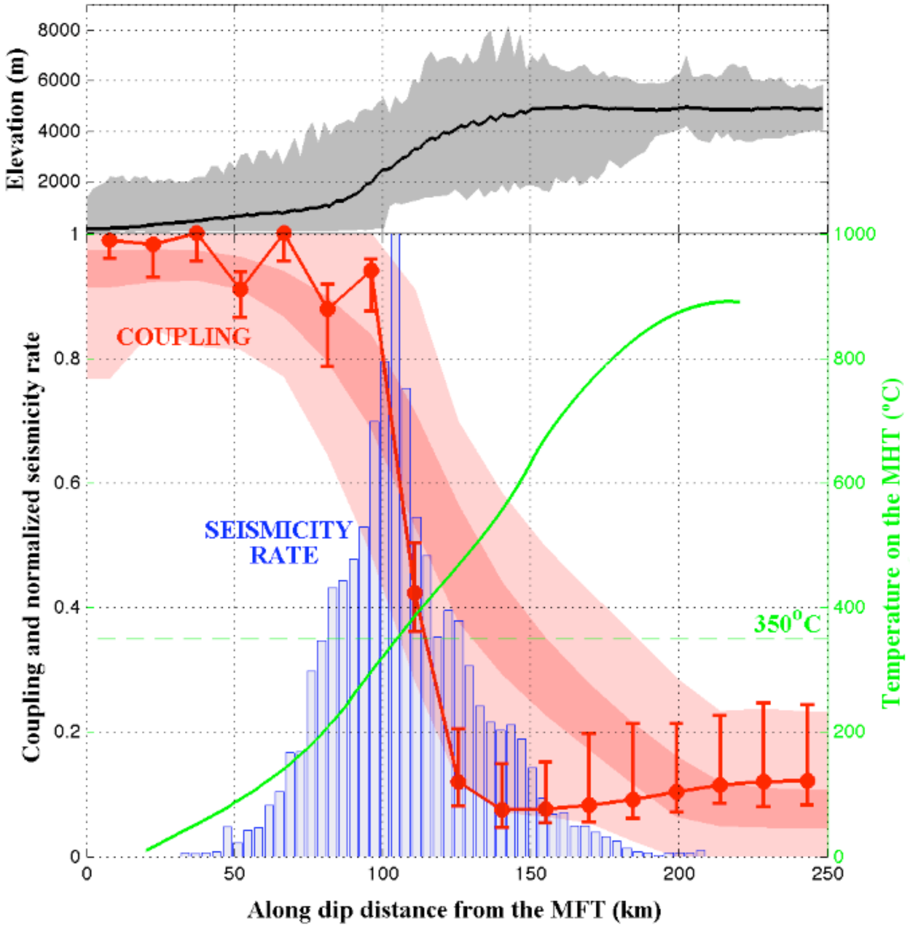


Figure 1.6: Upper plot: elevation profile. The black line represents the mean elevation, while the shaded grey area represents the whole elevation swath. Lower plot: Comparison between the coupling, temperature and seismicity rate along the dip direction. The red line with error bars corresponds to the coupling underneath the leveling line, where the resolution is the best. The shaded red curve in the background is a stack of the coupling on the whole fault, the darker red shaded area representing the 1-σ scatter of coupling, and the lighter red shaded area showing the whole scatter of coupling with respect to the distance to the MFT. The blue histogram shows the seismicity rate, normalized to a maximum value of 1. The green curve shows the temperature variation along a MHT dipping 10°, determined by *Herman et al.* (2010), corresponding to the duplex formation model. The thin dashed green line indicates the critical temperature of 350°C, above which frictional sliding is generally thought to be dominantly rate-strengthening, promoting stable sliding, according to lab experiments on quartzo-feldspathic rocks (*Blanpied et al.*, 1995; *Marone*, 1998).

advocated as an explanation for the seismicity cut-off generally observed at a depth of around 15km within continents. This explanation would hold well for the Himalaya as well as shown by the comparison of interseismic coupling with the thermal structure of the Himalaya of central Nepal, which is well constrained by thermokinematic modeling of thermochronological, thermometric and barometric data (*Herman et al.*, 2010). Here we use the best fitting thermal model which was determined from the inversion of the thermochronological thermobarometric data available for central Nepal. This model takes into account the accretion process that has resulted from the development of the Lesser Himalaya duplex over the last $\sim 10 - 15$Ma. Interseismic coupling is indeed observed to drop abruptly at the location where temperature increases from about 300°C to about 500°C according to the duplex model.

1.5.3 Relationship between geodetic strain and background seismicity

Previous studies had noticed that background seismicity along the Himalayan arc is clustered along a relatively narrow zone which follows the front of the high Himalaya (*Pandey et al.*, 1995, 1999) and which also approximately coincides with the downdip end of the locked fault zone (*Cattin and Avouac*, 2000; *Bollinger et al.*, 2004). This correlation suggests that seismicity is triggered by quasistatic stress build up in the interseismic period at the tip of the creeping zone (*Cattin and Avouac*, 2000; *Bollinger et al.*, 2004) . This interpretation is confirmed by our study: Figure 1.7 shows interseismic stress accumulation on the MHT derived from the interseismic coupling pattern together with the distribution of seismicity relocated with the double difference technique (*Waldhauser and Ellsworth*, 2000). The stress represented here corresponds to that on a planar fault where each patch would have a purely dip slip motion at a rate of 20×(1-ISC) mm/yr, assuming a shear modulus of 30 GPa and a Poisson ratio of 0.25. Such a rough approximation is justified by the fact that the stress rate estimation is limited anyways by the resolution of the inversion.

Clearly most of the microseismicity falls in the area of maximum Coulomb stress increase in the interseismic period. Only the along dip variations of stress rate should be regarded as a valuable information: along strike inhomogeneities in the stress pattern computed are here mostly due to inhomogeneities of the resolution, as discussed earlier. The stress accumulation rate is maximum underneath the leveling line because the locked-creeping transition is resolved with more accuracy there. Since this transition seems always sharper than what the resolution offers, the values of the stress accumulated plotted on Figure 1.7 should be seen as lower boundaries of the real values.

In fact, the seismicity does not occur on the MHT per se but rather within a 5-10

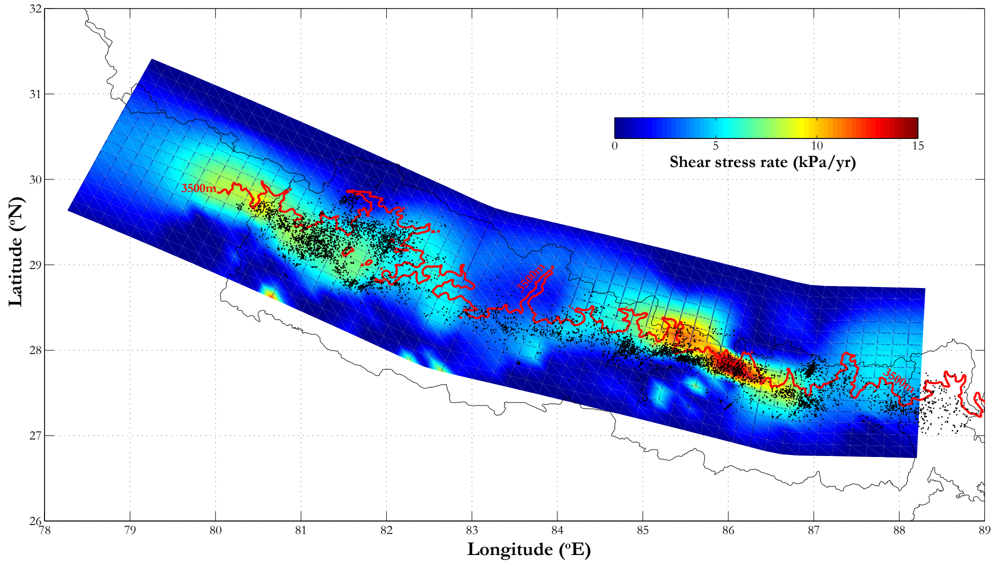


Figure 1.7: Mapview of the midcrustal microseismicity from 1996 to 2008 superposed to the map of the shear stress accumulation rate on the MHT, deduced from the coupling pattern. The thick red line represents the 3500m elevation contour line above which the seismicity seems to drop.

Catalog	Time span (yrs)	Moment released seismically (Nm/yr)	Fraction of moment accumulated[1] (%)
Historic[2]	~ 500	1.8×10^{19}	27.3
Historic[3]	~ 500	0.9×10^{19}	13.6
CMT	35	2.9×10^{18}	4.4
CMT micro[4]	35	3.4×10^{17}	0.5
NSC[5]	6	1.2×10^{17}	0.2

Table 1.2: Moment released during earthquakes annually according to different seismicity catalogs. This moment released is compared to the rate of moment deficit of 6.6×10^{19} Nm/yr inferred in this study.

kilometer size volume around the downdip end of the locked fault zone. The moment released by the background seismicity amounts to 1.2×10^{17} Nm/yr which represents less than 0.2% of the deficit of moment accumulating due to interseismic locking of the MHT (table 1.2). This quantity was estimated by converting local magnitudes reported in the catalogue of the National Seismological Centre (NSC) in Nepal from 1995 to 2001, a period of homogeneous completeness over the whole network, into moment magnitudes and by summing the scalar moments. To do so, the local magnitudes (M_L^{NSC}) reported in the

NSC catalogue were converted into moment magnitudes (M_W^{NSC}) from a subset of events reported also in the CMT calatogue (NSC, personal communication) yiedling:

$$M_W^{NSC} = 0.84 M_L^{NSC} + 0.21. \tag{1.7}$$

So background seismicity does not contribute much to releasing interseismic stress build up but it does reflect areas of most rapid stress increase.

Interseismic stress build up is probably not the only factor controlling the distribution of background seismicity as the seismicity is observed to shut off underneath the higher Himalaya, where the elevation gets higher than 3500m (Figure 1.7) (*Avouac*, 2003; *Bollinger et al.*, 2004). This correlation can be simply interpreted as the effect of the topography on the stress field: where the elevation is higher than 3500m the principal stress becomes vertical. As a result, optimally oriented faults correspond to normal fault planes whose rupture is actually inhibited by interseismic stress buildup (*Bollinger et al.*, 2004).

1.5.4 Moment deficit accumulation rate, return period and magnitude of the largest plausible earthquake

Given the pattern of interseismic coupling and the long term slip rate on the MHT derived from this study, locking of the MHT has resulted in the accumulation of a deficit of moment of about $\dot{\mathcal{M}}_0 = 6.6 \pm 0.4 \times 10^{19}$ Nm/yr (assuming a shear modulus of 30 GPa) over the last 20 years covered by the dataset analyzed in this study. This value is quite robust with respect to the parameters of the inversion, for instance the Laplacian smoothing (Figure 1.8(a)) or the direction selected for the extension of the Tibetan plateau (Figure S5).

Over the longer run, this deficit of moment has to be compensated by transient slip events along the MHT: this must presumably be the result of large Himalayan earthquakes and associated afterslip. The rate of accumulation of moment deficit can therefore be used to estimate the return period of large earthquakes.

The return period estimate is related to a number of additional parameters that are not necessarily available, and some assumptions hence have to be made. First, we assume that the rate of moment deficit accumulation can be extrapolated over the whole duration of the interseismic period. This assumption is justified by the lack of evidence for any significant temporal change over the period analyzed here (i.e., 13 years for the stations GUMB and DAMA, see the time series at DAMA on supplementary Figure S3), and for the insignificant difference between eastern Nepal, which last produced a large ($M_W \sim 8.1$, *Ambraseys and Douglas* (2004)) earthquake in 1934, and western Nepal where no large earthquake has occurred since 1505. Moreover, both the convergence velocities in eastern and western Nepal approximately match the geological slip rate across the MFT proposed by *Lavé and Avouac* (2000) within uncertainties. Since that geological slip rate encompasses several earthquake cycles, it represents the average convergence velocity over time, and is therefore the value to use when evaluating the moment deficit. Whether this value varies with time and today matches its mean value or is constant with time actually does not matter for our estimate of the moment deficit, as long as we use a value close enough to the average convergence rate, which is the case here. Time variations of the coupling pattern on the MHT also seem like a dubious eventuality, since such variations would have no reason to be uniform in space and would thus most likely engender lateral variations on the snapshot of coupling pattern that we observe today. As was mentioned earlier, such lateral variations do not show up in the present study. It is hence hard to imagine time variations of the coupling pattern large enough to modify the moment deficit accumulation rate by more than a few percents.

We also assume that a fraction α of the moment deficit is released by seismic slip through a distribution of earthquakes following a Gutenberg-Richter law (*Gutenberg and Richter*, 1954) up to a maximum magnitude corresponding to a moment $\mathcal{M}_{max}$, above which the seismicity rate drops to zero. The remaining fraction, $1 - \alpha$, is assumed to be released elastically by transient aseismic slip event (slow slip events, hereafter SSEs) or afterslip following large earthquakes. As was mentioned earlier, anelastic deformation of the crust is ruled out by the observation that the geological slip rate on the MFT is comparable to the shortening rate across the Himalayan range.

The oldest GPS stations in Nepal have now been recording daily positions for almost 13 years and no SSE has been identified in the time series (see for instance the time series of the station DAMA in supplementary Figure S3(b)). Known SSEs usually have return periods of less than a few years (e.g., review by *Schwartz and Rokosky*, 2007). This might be an artifact of the short observation time span of SSEs, which could not be detected until a couple decades ago owing to the lack of adequate instruments, so nothing rigorously bans a scenario of large unfrequent SSEs in Nepal releasing a major portion of the moment deficit. However, in absence of direct evidence for SSEs over the 20 years period covered by geodetic data, we assume in the following that SSE do not contribute significantly to the release of interseismic strain. This is a strong hypothesis that should be kept in mind hereafter. Afterslip generally tapers off within a year following the mainshock and can typically release around 25% of co-seismic slip as has been observed for intracontinental earthquakes (*Hsu et al.*, 2009b; *Perfettini and Avouac*, 2007) and a number of subduction zone earthquakes (*Melbourne et al.*, 2002; *Hsu and Bürgmann*, 2006; *Chlieh et al.*, 2008; *Perfettini et al.*, 2010). A few exceptions should be mentioned though. The afterslip of the Sanriku-Haruka-Oki earthquake (a typical interplate thrust event of moment magnitude $M_W = 7.6$) released an energy equivalent to a $M_W = 7.7$ earthquake (*Heki et al.*, 1997) ($\alpha \sim 0.4$), while the large afterslip following the 2004 $M_W = 6$ Parkfield earthquake would lead to a value of α as low as 0.25 (*Freed*, 2007). Based on relatively sparse data, it seems that the afterslip of the 2005 Kashmir earthquake released a moment significantly large in proportion of the co-seismic moment, corresponding to an α value ranging between 0.6 and 0.7 (*Jouanne et al.*, 2011). So a reasonable range of values for α is probably between 0.5 and 0.9, with a more probable value around 0.8.

Under those assumptions the recurrence time of earthquakes of moment $\mathcal{M}$ is (*Molnar*, 1979)

$$T(\mathcal{M}) = \frac{1}{1 - 2b/3} \frac{\mathcal{M}_{max}}{\alpha \dot{\mathcal{M}}_0} \left(\frac{\mathcal{M}}{\mathcal{M}_{max}} \right)^{2b/3}, \tag{1.8}$$

where the b-value of the Gutenberg-Richter distribution is usually close to 1. For $b = 1$, which is approximately the case for the crustal seismicity in Nepal (Figure 1.8(b)) the return period of the largest possible earthquakes on the MHT becomes

$$T_{b=1}(\mathcal{M}_{max}) = \frac{3\mathcal{M}_{max}}{\alpha \dot{\mathcal{M}}_0}. \tag{1.9}$$

Figure 1.8(b) displays a comparative Gutenberg-Richter plot of the different seismicity catalogs available in Nepal and equation (1.9) for $\alpha = 1$, and $\mathcal{M}_{max} = 8$, 9 and 10. Three catalogs are represented on this plot. The first one is the microseismicity monitored between 1995 and 2001 by the National Seismological Centre (NSC) in Nepal. 1995-2001 corresponds to the period where the seismic network was functioning well enough to have a homogeneous magnitude completeness on the whole Nepalese territory.

The second catalog is the CMT catalog, that covers the last 35 years, and for which we have selected earthquakes with a dip-slip focal mechanism (rake $= 90° \pm 45°$) corresponding in map view to the midcrustal cluster. Here we have considered the whole Himalayan arc, over its full length extent (almost 3000 km), and rescaled to the territory of Nepal which extends over about a third of the full length of the Himalayan arc. One should keep in mind that this catalog is largely dominated by the 2005 M_W 7.6 Kashmir earthquake and its aftershocks (45 out of the 69 events of the whole catalog). Finally the historic catalog

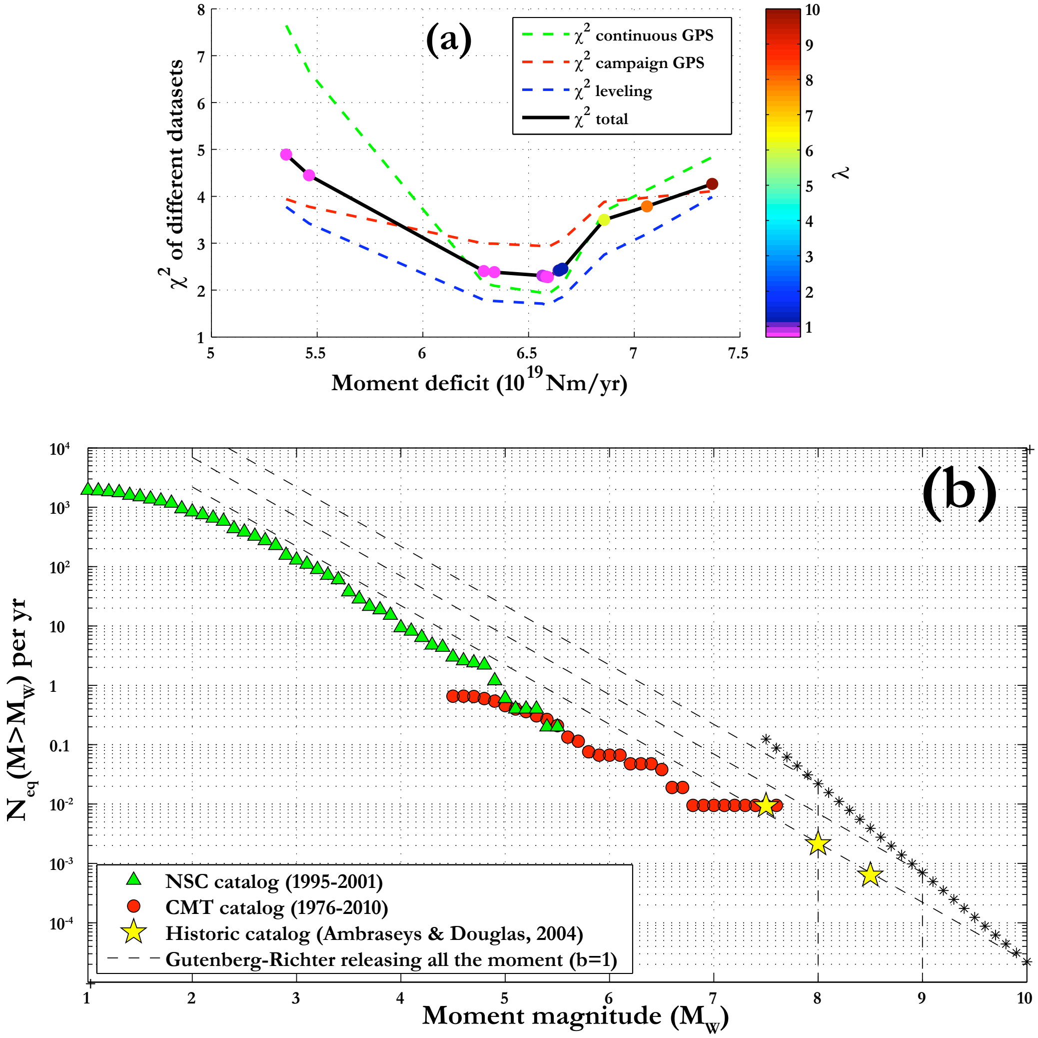


Figure 1.8: Seismic hazard evaluation. (a) Impact of the weights applied to the Laplacian (parameter λ in equation (1.6)). The plots shows the value of the reduced χ^2 of the fit as well as the moment deficit accumulation rate for each inversion. (b) Gutenberg-Richter plot of the seismicity in Nepal, using the different catalogs available: The NSC catalog (1995-2001), the CMT catalog (1976 - 2010) and an historic catalog compiled using the catalog from *Ambraseys and Douglas* (2004). We used the last 500 years of the historic catalog for $M_W > 8$ earthquakes, and the last 200 years for $M_W > 7.5$ earthquakes. The dotted lines are the distribution that the seismicity should follow if 100% of the moment deficit was released seismically following a Gutenberg-Richter distribution with $b = 1$, up to a given maximum magnitude of 8, 9 and 10. The asterisk line shows, for a given maximum possible magnitude for Himalayan earthquakes, the return period of such earthquakes.

has been compiled using the catalog from *Ambraseys and Douglas* (2004), accounting for all earthquakes over the whole Himalayan arc, considering again that Nepal covers one third of the chain. The only $M_W \geq 8.5$ earthquake of this catalog in the past 500 years is the 1950

Assam earthquake. Over the same period of time, it has 3 $M_W \geq 8$ earthquakes (1505, 1934, 1950), and in the past 200 years, 6 $M_W \geq 7.5$ earthquakes.

Above their magnitude of completeness, the earthquakes catalogs are fairly consistent with $b = 1$. However, they fall noticeably below the 3 dotted lines corresponding to equation (1.9), meaning that they do not account for all the moment deficit that accumulates in the interseismic period. Clearly the known historical and instrumental seismicity falls well short of balancing interseismic strain buildup. If we assume that earthquakes in Nepal never exceed a moment magnitude of 8.5, and that every large earthquake is documented within the catalogs, seismicity over the last 500 years would account for less than 20% of slip deficit due to locking of the MHT in the interseismic period (table 1.2). The contribution is even less if we assume a lower possible magnitude (say $M_W = 8$). This mismatch, is too large to be due to the contribution of afterslip. If we now assume that the largest earthquakes on the MHT could reach a magnitude higher than the $M_W \sim 8.5$ magnitude estimated for the 1950 Assam event, the mismatch is reduced. We note that the historic seismicity also seems to line up in favor of the occurrence of very large ($M_W > 9$) earthquakes in Nepal. Such earthquakes would have a very long return period: a maximum magnitude $M_W^{\max} = 9.2$ would have a return period of the order of 3000 years if all the moment deficit was released seismically (see supplementary Figure S6(a)). In the absence of a clear segmentation of interseismic stress build up along the Nepal Himalaya, we cannot exclude the possibility of such a large event. As a comparison, it is interesting to note that western Nepal has not apparently ruptured since the 1505 earthquake (*Ambraseys and Douglas*, 2004). If the moment deficit accumulation has proceeded at a constant rate since then, a release now of this moment deficit between the 1934 Bihar earthquake and the western border of Nepal (i.e., about 500 km of fault length) could generate up to a M_W 8.9 earthquake.

Conclusion

The denser network of geodetic data in Nepal brings better kinematic constraints on the convergence of India underneath the Tibetan plateau. The MHT appears to be nearly fully locked from the surface to beneath the front of the high Himalaya, over a width of about 100km. Interseismic coupling decreases abruptly, within a transition zone probably narrower than 30km. This transition occurs at a depth of about 15-20 km, where the temperature on the MHT is estimated to reach 350°C. This might reflect that stable aseismic sliding is promoted where the temperature exceeds 350°C as inferred from laboratory experiments and observations in other continental contexts (*Blanpied et al.*, 1995; *Marone*, 1998; *Hsu et al.*, 2009b). This favors the scenario of a primary control by temperature of the locked-creeping transition, similar to the conclusions drawn by *Hsu et al.* (2009b) on the Chelungpu fault in Taiwan and by *Brooks et al.* (2011) on the Mandeyapecua thrust fault in the central Andean backarc. The microseismicity on the MHT seems to cluster where the shear stress accumulation is the greatest, and drops under topography greater than 3500m of elevation, i.e., where the principal Coulomb stresses become vertical. The apparent segmentation of the microseismicity then comes off as a result of the competition between the relative positions of the 3500m contour line and of the locked-creeping transition, where the stress rate is the greatest. The lack of any apparent lateral variation of coupling is an interesting result, since it differs from observations at subduction zones, whose patterns of coupling exhibit noticeable segmentations ((*Chlieh et al.*, 2008; *Moreno et al.*, 2010; *Suwa et al.*, 2006; *Ozawa et al.*, 2011; *Loveless and Meade*, 2010; *Freymueller et al.*, 2000; *Wallace et al.*, 2004). This might point to a fundamental difference between intracontinental and subduction megathrust. In any case, the rate of accumulation of moment deficit on the MHT within Nepal is large ($6.6 \pm 0.4 \times 10^{19}$ Nm/yr), and comparison with the historical seismicity suggests that infrequent (with return period larger than 1000yr) events with magnitude larger than the $M_W \sim 8$ value assigned to the largest known earthquakes of 1934 and 1505 should be taken into consideration, as inferences based on paleoseismological investigations have also suggested (*Lavé et al.*, 2005). However, one should keep in mind that those seismic hazard assessment rely on a few hypothesis (no significant release of moment by afterslip or slow slip events) that could alter our conclusions if proven inexact.

Acknowledgments

First of all, this project was supported by the Gordon and Betty Moore Foundation, through the Tectonics Observatory, NSF grant EAR#0838495 and the China Geological Survey (1212011121265). We would like to thank reviewers Michael Taylor from the University of Kansas and Rebecca Bendick from the University of Montana for useful comments. We thank the staff of the NSC in Nepal for their hard work on the seismic network and the continuous GPS stations, and for providing us with the NSC seismic catalog for this study. We thank Michel Dandine and Dili Ram Tiwari for their contribution to the seismicity catalog of Nepal. We also thank Frederic Herman for providing values of his temperature models. This is Tectonic Observatory's contribution number 193.

Coupling on the MHT - Supplementary material

S.1 Map of the GPS network

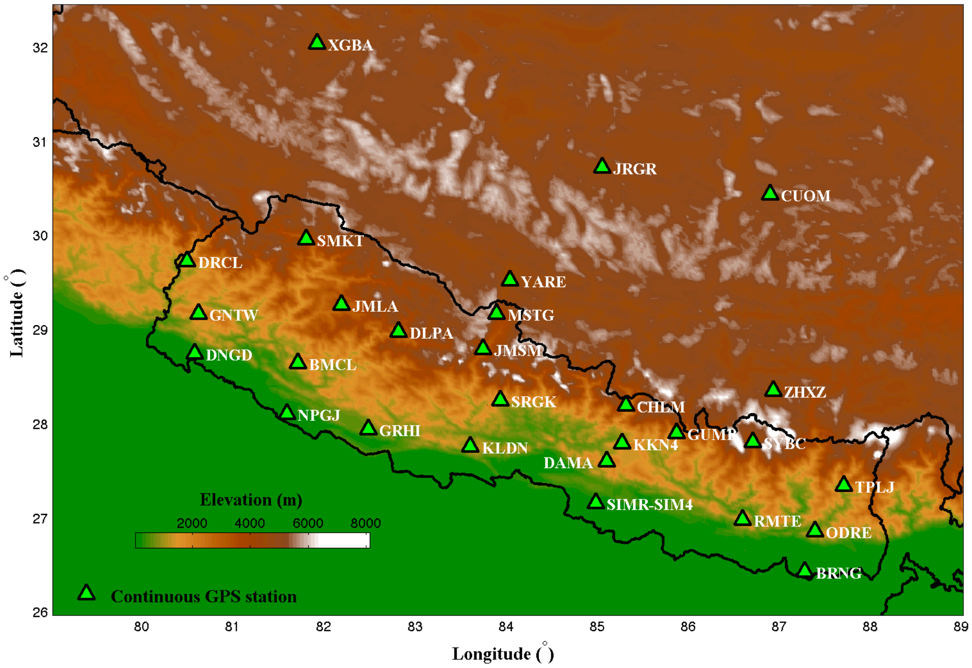


Figure S1: Map indicating the names and position of the continuous GPS stations used in this study to determine the pattern of coupling on the MHT.

S.2 Derivation of secular velocities from the GPS time series

S.2.1 Noise model for the inversion

Several studies of GPS time series have established that the daily estimates of GPS positions are temporally correlated (*Langbein and Johnson*, 1997; *Zhang et al.*, 1997; *Mao et al.*, 1999; *Williams*, 2003a; *Williams et al.*, 2004). Assuming a purely white noise model is therefore incorrect and although it doesn't affect much the value of the final parameters inverted for, it results in a dramatic underestimation of their uncertainties. We thus add to the white noise in our GPS time series a component of colored noise, i.e. a noise that has a power spectrum of the form:

$$P_\kappa(f) \propto f^\kappa, \tag{S1}$$

where f is the temporal frequency, and κ is called the spectral index (*Mandelbrot and Van Ness*, 1968). The spectral index is estimated for each time series by first fitting equation (1.1) (see main paper) to the time series assuming a white noise, and computing a periodogram of the fit's residuals. The spectral index is then estimated by fitting to the power spectrum a combination of white and colored noise (figure S2):

$$P(f) = P_0 + P_c f^\kappa, \tag{S2}$$

where P_0 and P_c are the respective amplitudes of the white and the colored noise.

Once the spectral index is estimated, we build the covariance matrix of the data as the sum of white and colored noise covariance matrices. The relative amplitudes of both noises are estimated by a Maximum Likelihood Estimation (MLE) method (*Williams et al.*, 2004).

The covariance matrix for the white noise is the usual diagonal covariance matrix

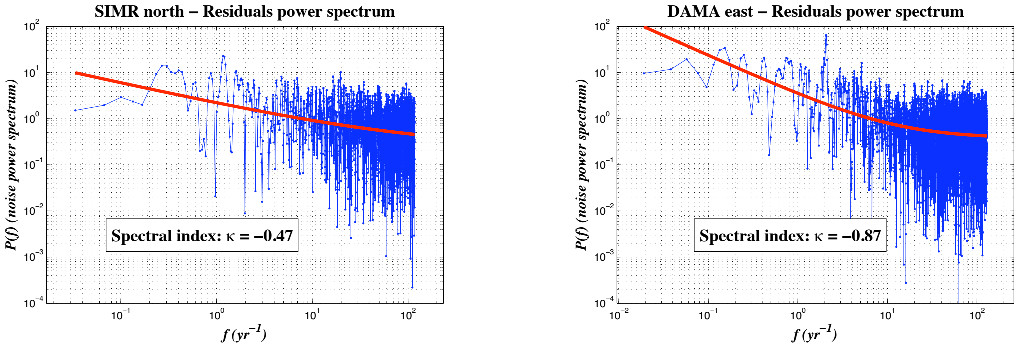
$$\mathbf{C_w} = \mathrm{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2),$$

where σ_i is the standard deviation of data point number i. The colored noise covariance matrix $\mathbf{C}_\kappa$ is built following an adaptation of the method described in *Williams* (2003a):

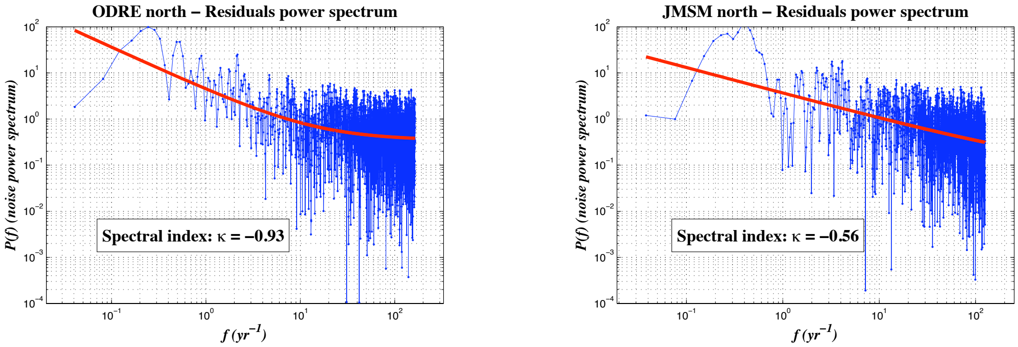
$$\mathbf{C}_\kappa = \Delta t_s \mathbf{T}\mathbf{T}^T, \tag{S3}$$

where Δt_s is the sampling interval (so $\Delta t_s = 1$ day for GPS time series), and the matrix $\mathbf{T}$ is defined as:

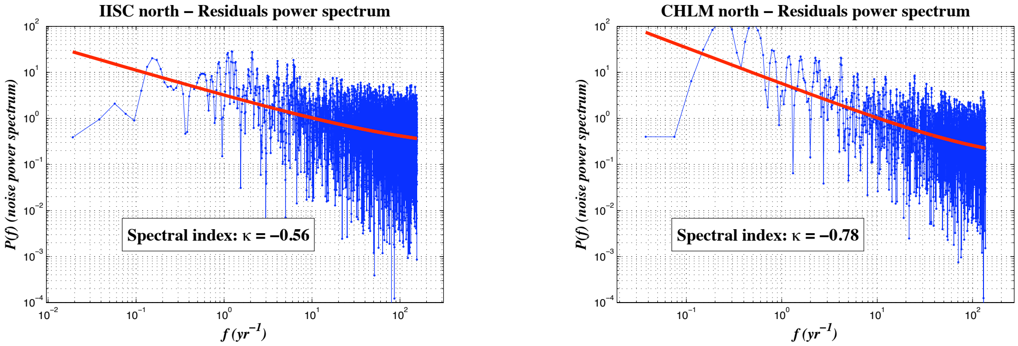
$$\mathbf{T} = \begin{pmatrix} \psi_0 & 0 & 0 & \ldots & 0 \\ \psi_1 & \psi_0 & 0 & \ldots & 0 \\ \psi_2 & \psi_1 & \psi_0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \psi_{n-1} & \psi_{n-2} & \psi_{n-3} & \ldots & \psi_0 \end{pmatrix}, \tag{S4}$$



(a) SIMR, north component (b) DAMA, east component



(c) ODRE, north component (d) JMSM, north component



(e) IISC, north component (f) CHLM, north component

Figure S2: Power spectra of the noise (blue curve) for a sample of time series and fit assuming a combination of white noise and colored noise (red curve). The spectral index κ of the colored noise is indicated on each plot.

where the quantities ψ_n are defined by the recurrence formula:

$$\begin{cases} \psi_0 = 1 \\ \psi_{n+1} = \frac{n-\kappa/2}{n+1}\ \psi_n \end{cases} .$$

The rows and columns corresponding to times with no data are then removed from the covariance matrix.

The final data covariance matrix is given by

$$\mathbf{C_D} = a^2\mathbf{C_w} + b^2\mathbf{C}_\kappa, . \tag{S5}$$

where a and b are the parameters to be estimated by MLE, measuring respectively the amplitude of white and colored noise. Assuming a Gaussian distribution of the uncertainties on GPS positions, the likelihood that has to be maximized with respect to a and b is then

$$\text{likelihood}(\mathbf{C_D}) = \frac{1}{(2\pi)^{N/2}(\det \mathbf{C_D})^{1/2}} e^{-\frac{1}{2}\mathbf{r}^T\mathbf{C_D}^{-1}\mathbf{r}}, \tag{S6}$$

where $\mathbf{r}$ is vector of residuals of the fit and N is the number of daily GPS positions available.

The fit and residuals on some time series are shown on figure S3.

S.2.2 Uncertainties due to unmodeled steps in the time series

Steps in the time series can be of many different origins, being actually tectonic, environmental or coming from equipment malfunction, human error, etc. (*Williams*, 2003b). The ones large enough to be detected are included in the model (equation (1.1)), but smaller ones remain unnoticed and affect the estimates of model parameters and their uncertainties. Therefore, those uncertainties have to be adjusted accordingly. For convenience, we will assume that those unmodeled steps account for all the errors on the model.

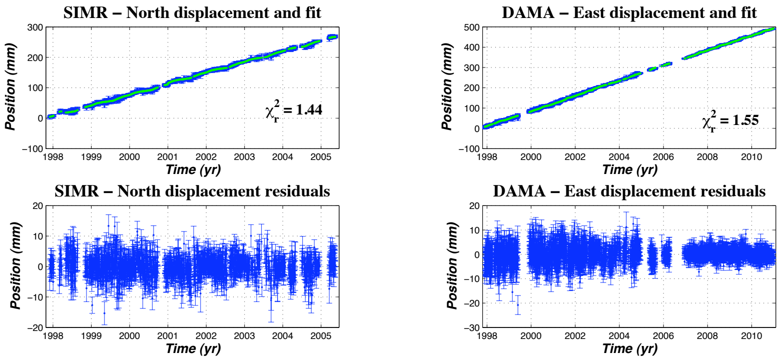
Those steps are assumed to happen at a frequency ν, and to have a random Gaussian amplitude $\mathcal{N}\left(0, \sigma_x^2\right)$. The standard deviation on the secular velocity due to those steps is then (*Williams*, 2003b)

$$\sigma_v = \frac{\sigma_x\sqrt{\nu}}{\sqrt{T}}, \tag{S7}$$

where T is the length of the time series.

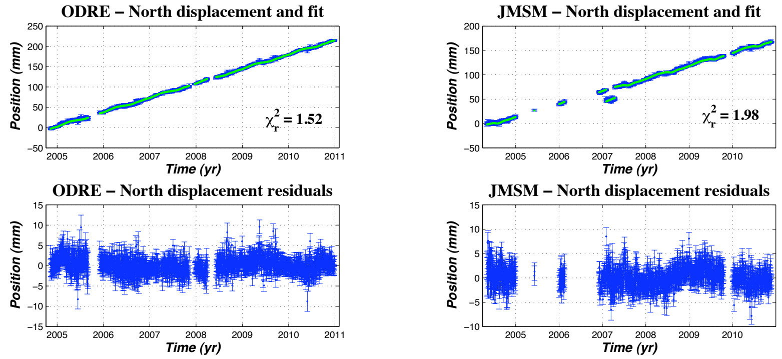
In the case of our GPS time series, the amplitude of the steps that were actually detected was always greater than 1.5 time the median value of the uncertainties on the daily positions in the time series. We hence take $\sigma_x = \langle\sigma_D\rangle$, where $\langle . \rangle$ denotes the median value and σ_D is the uncertainty on daily positions of the time series.

We estimate ν through the following considerations. First, the steps that were large enough to be detected in the time series happened on average once every 5 years. Assuming that the smaller the steps are, the more frequent they would be, the value for ν should be



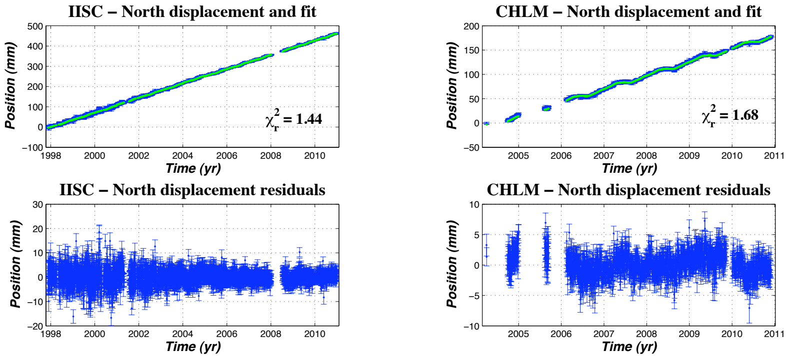
(a) SIMR, north component

(b) DAMA, east component



(c) ODRE, north component

(d) JMSM, north component



(e) IISC, north component

(f) CHLM, north component

Figure S3: Fits and residuals of the continuous GPS time series at some stations. For each time series, the upper plot shows the raw data (blue curve) and the fit (green curve) with equation (1.1). The value of the reduced chi square χ_r^2 of each fit is indicated on the upper plot's lower right corner of the corresponding figure.

greater than 1/5. On the other hand, a value of ν overestimated ($\nu > 1$ in this case) results in larger uncertainties on the secular velocity, and eventually leads to values of a reduced chi square smaller than 1 when one fits the Euler pole of the Indian plate in the ITRF 2005 reference frame (see section 1.3.3), indicating that the uncertainties on the GPS velocities are probably overestimated. As a result, we chose a value of $\nu = 1/3$, which gives the final formula for the uncertainties on the secular velocity due to unmodeled steps:

$$\sigma_v = \frac{\langle \sigma_D \rangle}{\sqrt{3T}}. \tag{S8}$$

The velocities and corresponding uncertainties that we obtain at the GPS stations used in this paper as well as at the DORIS stations COLA and EVEB are given in table S2.

S.3 Slip resolution and Laplacian

The result of our inversion should be assessed in view of its resolution. This information is contained in the resolution matrix:

$$\mathbf{R} = \left(\mathbf{G}^T \mathbf{C_d}^{-1} \mathbf{G} + \mathbf{\Lambda}^T \mathbf{\Lambda}\right)^{-1} \mathbf{G}^T \mathbf{C_d}^{-1} \mathbf{G}, \tag{S9}$$

where **G** is the Green's matrix defined in equation (1.5) from the main paper, $\mathbf{C_d}$ is the data covariance matrix and $\mathbf{\Lambda}$ is the Laplacian matrix. The diagonal of **R** tells how well the slip value on each patch can be retrieved by the inversion. However, it doesn't express how each patch correlates with its neighbors. This information is contained in each of the individual columns of **R**: column number i is the vector of parameters (i.e. the slip on each patch) returned by the inversion from an input dataset corresponding to a unit slip on patch i and no slip on other patches. Usually, what the inversion returns is slip on a more

Station ID	Site name	Latitude (°N)	Longitude (°E)	Elevation (m)
BAN2	Bangalore	13.03431	77.51161	832
DGAR	Diego Garcia	-7.26968	72.37024	-65
GUAO	Guao	43.47111	87.17731	2029
GUAM	Guam Observatory	13.58933	144.86836	202
HYDE	Hyderabad	17.41726	78.55087	442
IISC	Indian Inst. Science	13.02117	77.57038	844
KUNM	Kunming	25.02954	102.79712	1986
LHAS	Lhasa	29.65734	91.10399	3625
LHAZ	Lhasa2	29.67533	91.10403	3625
POL2	Poligan IVTAN 2	42.67977	74.69427	1714
SELE	Selezaschita	43.17873	77.01690	1342
TAIW	Taipei	25.02133	121.53654	44
URUM	Urumqi	43.80795	87.60067	859
WUHN	Wuhan	30.53165	114.35726	26

Table S1: List of IGS sites included in the daily regional processing.

			Velocities in ITRF05 (mm/yr)			Time of operation	
Station	lon (°E)	lat (°N)	V_e	V_n	V_u	Init.	End
DAMA	85.1077	27.6081	36.73 ± 0.45	34.21 ± 0.33	1.46 ± 1.37	Nov 1997	current
GUMB	85.8775	27.9098	35.88 ± 0.44	28.2 ± 0.34	5.7 ± 1.19	Nov 1997	current
SIMR	84.9844	27.1646	37.05 ± 1.03	34.82 ± 0.58	2.68 ± 2.72	Nov 1997	Apr. 2005
BRNG	87.2813	26.4387	37.7 ± 2.83	36.15 ± 3.49	−2.2 ± 3.55	Mar 2004	May 2009
BRN2	87.272	26.5197	38.46 ± 1.34	33.61 ± 1.82	8.13 ± 5.48	May 2009	current
CHLM	85.3154	28.2054	36.65 ± 0.48	27.55 ± 0.35	4.36 ± 1.2	Mar 2004	current
JMSM	83.7467	28.8044	34.33 ± 0.54	26.28 ± 0.39	3.28 ± 1.36	Ma. 2004	current
KKN4	85.2788	27.8008	36.12 ± 0.45	32.27 ± 0.4	1.13 ± 1.23	Jan 2004	current
KLDN	83.6119	27.7669	35.66 ± 0.43	34.64 ± 0.36	1.69 ± 1.17	Apr 2004	current
MSTG	83.8946	29.1789	34.68 ± 1.3	24.19 ± 1.05	5.79 ± 3.7	Apr 2004	Sept 2004[a]
MST2	83.953	29.1778	31.64 ± 1.59	23.73 ± 1.13	1.81 ± 4.09	Oct 2009	current
ODRE	87.3921	26.8662	38.71 ± 0.52	35.44 ± 0.37	−3.02 ± 1.5	Mar 2004	current
SIM4	84.99	27.17	37.13 ± 0.63	35.32 ± 0.5	−1.4 ± 1.8	Mar 2004	current
SRGK	83.9358	28.2603	35.62 ± 1.03	30.65 ± 0.9	4.52 ± 3.01	Mar 2005	Feb 2007
TPLJ	87.71	27.35	37.98 ± 0.55	31.24 ± 0.32	1.25 ± 1.03	Mar 2004	current
BMCL	81.7144	28.6558	34.51 ± 0.93	33.5 ± 0.47	0.47 ± 2.29	Mar 2007	current
DLPA	82.8204	28.9853	34.85 ± 0.63	25 ± 0.51	1.23 ± 1.68	May 2007	current
GRHI	82.4914	27.9509	35.09 ± 0.62	32.27 ± 0.54	3.97 ± 1.64	May 2007	current
JMLA	82.1923	29.2742	32.04 ± 0.76	26.14 ± 0.45	2.1 ± 1.6	May 2007	current
NPGJ	81.5953	28.1172	35.22 ± 0.67	39.04 ± 0.74	−0.16 ± 1.63	May 2007	current
BYNA	81.2007	29.4742	31.5 ± 1.38	26 ± 0.74	1.69 ± 2.39	May 2008	current
DNGD	80.5818	28.7545	35.29 ± 0.67	30.14 ± 0.85	−0.48 ± 1.8	May 2008	current
DRCL	80.5009	29.7338	31.41 ± 0.72	29.78 ± 1.2	2.64 ± 2.34	Mar 2008	current
GNTW	80.6262	29.1765	33.57 ± 0.67	33.68 ± 0.63	0.72 ± 3.15	Apr 2008	current
RMJT	86.55	27.3051	35.08 ± 1.62	32.58 ± 1.48	−1.04 ± 4.36	Oct 2008	current
RMTE	86.5971	26.991	35.86 ± 0.78	33.49 ± 0.69	1.44 ± 2.02	Sep 2008	current
SMKT	81.8065	29.9694	29.82 ± 0.71	21.84 ± 0.52	3.73 ± 1.83	May 2008	current
SYBC	86.7125	27.8142	35.55 ± 1.18	25.93 ± 0.96	7.14 ± 2.88	Oct 2008	current
CUOM	86.9039	30.4451	40.52 ± 0.59	20.38 ± 0.52	0.3 ± 2.05	Oct 2006	current
JRGR	85.0568	30.7286	36.11 ± 0.59	20.39 ± 0.5	3.03 ± 1.57	Mar 2007	current
XGBA	81.9259	32.0469	29.19 ± 1.55	18.25 ± 1.43	1.47 ± 4.25	Mar 2007	Sep 2007[b]
YARE	84.0431	29.5344	35.56 ± 0.68	24.14 ± 0.53	3.56 ± 1.86	Oct 2006	current
ZHXZ	86.9396	28.3569	37.86 ± 0.54	24.36 ± 0.45	1.87 ± 1.88	Oct 2006	current
MALD	73.526	4.189	43.35 ± 0.65	34.9 ± 0.49	−5.1 ± 1.93	Jul 1999	May 2006
HYDE	78.551	17.417	39.24 ± 0.49	35.24 ± 0.36	0.48 ± 1.18	Sept 2002	current
IISC	77.5704	13.0212	41.74 ± 0.47	35.06 ± 0.35	−0.05 ± 1.26	Oct 1997	current
COLA	79.8741	6.892	44.54 ± 4.8	35.33 ± 2.99	0.44 ± 3.78	Jan 1993	Sep 2004
EVEB	86.8131	27.9581	37.08 ± 4.4	25.32 ± 3.17	2.06 ± 3.64	May 1993	current

Table S2: Estimates of the secular velocity at the continuous GPS stations in ITRF 2005 and dates of operation of each station. The uncertainties on the velocities indicated are the 1-σ uncertainties. See text for details on the derivation of those quantities. Gaps in the time series are not unfrequent, and one should keep in mind that they are not indicated in this table.

---

[a] A 2-day campaign measurement has also been done with a different antenna on the station's monument in October 2009.

[b] 4 additional points in May 2009 made the positions at this station exploitable.

or less spread area centered on patch i. The characteristic size of this area is estimated by fitting a bell curve to the slip on the patches as a function of distance to patch i (*Lohman*, 2004), and taking the standard deviation of that bell curve. Namely, for each patch i, we find the distance w_i that minimizes the quantity:

$$\chi_i^2 = \sum_{j=1}^{N_\mathrm{p}} \left(\left| \frac{R_{ji}}{R_{ii}} \right| - e^{-\frac{d_{ij}^2}{2w_i^2}} \right)^2 , \tag{S10}$$

where N_p is the number of patches on the fault, R_{ji} is the value of the coefficient (j, i) of the resolution matrix $\mathbf{R}$ (row j and column i), and d_{ij} is the distance between patches i and j.

This idea of an estimate of the resolution scale on each patch is also used in order to more efficiently smooth our model by weighting the Laplacian according to the resolution on each patch. Since the Laplacian matrix is not yet available (this is what we try to determine), we compute a first resolution matrix using the Moore-Penrose pseudoinverse matrix (*Aster et al.*, 2005), keeping only the singular values larger than 10% of the maximum one. We then compute how far each patch correlates with its neighbors with the method previously described applied to this resolution matrix. Finally, each line of the Laplacian matrix is weighted by the decimal logarithm of the resolution size on the corresponding patch.

S.4 Supplementary figures on the pattern of coupling on the MHT

S.4.1 Laplacian smoothing

On figure S4 we test how different values of the Laplacian smoothing affect the estimate of the moment deficit accumulated every year. Weights assigned to the Laplacian too small ($\lambda < 0.8$) lead to models featuring locked patches only underneath data points, right next to creeping patches. Besides being unphysical and resulting in very high reduced chi squares, such models are highly dependent on the data spatial distribution and must then be rejected. A smoothing too large ($\lambda > 5$) tends to lead to a fault locked further at depth, and with a very smooth locked-creeping transition, which doesn't fit the data anymore (reduced $\chi^2 > 3$ on figure S4). Within the range of Laplacian weight $0.8 < \lambda < 5$, the moment deficit accumulated each year remains within the uncertainties determined by the inversion.

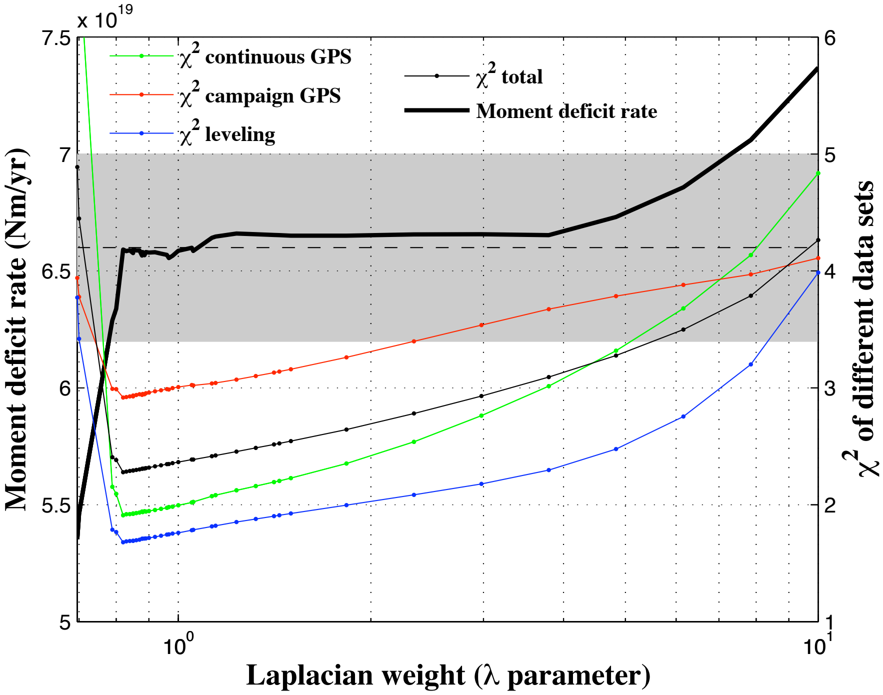


Figure S4: Variation of the χ^2 of the fit and of the moment deficit rate for different values of the weight attributed to the Laplacian in the inversion. The black curve shows the moment deficit accumulated every year as a function of the weight attributed to the Laplacian. The dashed black line and grey shaded area represent the rate of moment deficit with uncertainties derived in this study, i.e $\dot{\mathcal{M}}_0 = 6.6 \pm 0.4 \times 10^{19}$ Nm/yr. The green, red and blue curves respectively represent the value of the χ^2 of the fit to the continuous GPS, campaign and leveling data.

S.4.2 Direction of extension of the Tibetan plateau

Figure S5 shows the sensitivity of the long term velocity and the moment accumulation rate estimated in this study to the direction chosen for the extension of the Tibetan Plateau. The direction N98.2E has been chosen because it is the one that affects the least the estimates of the long term velocities (it is the 'most perpendicular' direction to those velocities, i.e. it is the direction onto which the sum of the projections of the East and West long term velocities reaches a minimum). But there is no real reason to prevent this direction from varying by a few degrees from the N98.2E azimuth. Figure S5 shows that even by changing this direction by 10°, the final values of the parameters remain within their estimated uncertainties.

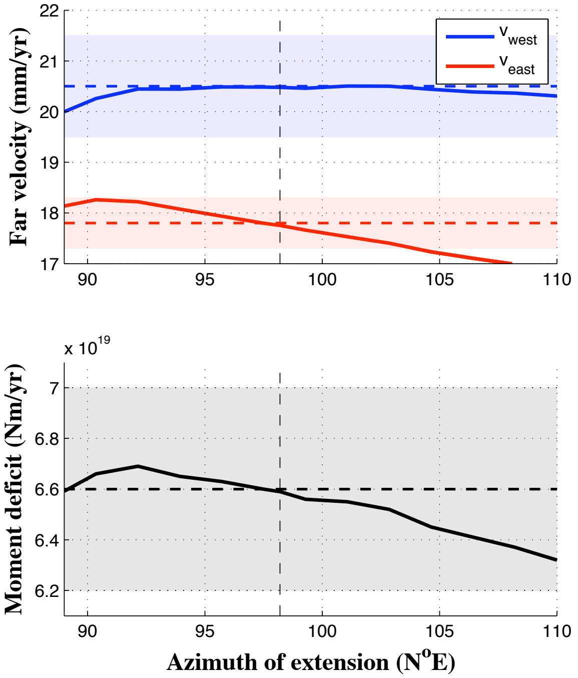
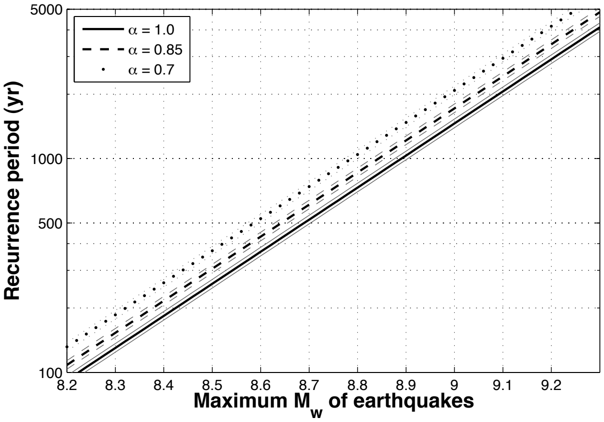


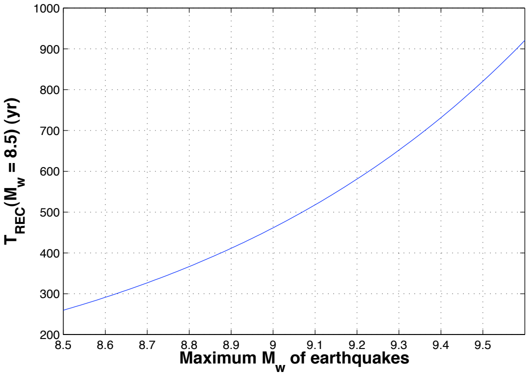
Figure S5: Impact of the azimuth selected for the extension of the Tibetan plateau on the long term East and West velocities (upper plot) and the moment deficit rate (lower plot). The solid thick lines represent the values of the parameters with respect the azimuth, the horizontal dashed lines and filled area of corresponding colors are the values with 1-σ uncertainties that we retained in this study (corresponding to an azimuth of N98.2E): $V_e = 17.8 \pm 0.5$ mm/yr, $V_w = 20.5 \pm 1$ mm/yr and $\dot{\mathcal{M}}_0 = 6.6 \pm 0.4 \times 10^{19}$ Nm/yr.

S.4.3 Recurrence time of large eartquakes

Assuming that the moment deficit of $\dot{\mathcal{M}}_0 = 6.6 \pm 0.4 \times 10^{19}$ Nm/yr computed in the main paper was released through earthquakes following a Gutenberg-Richter distribution up to a maximum magnitude, the recurrence time of those largest earthquakes (corresponding to those largest magnitudes) is plotted on figure S6(a). The black lines (solid, dashed and dotted) correspond to $\dot{\mathcal{M}}_0 = 6.6 \times 10^{19}$ Nm/yr, with different percentages of this moment deficit being released seismically, while the grey surrounding lines show the extent corresponding to the uncertainties on $\dot{\mathcal{M}}_0$. This plot shows that earthquakes as large as the 1950 Assam earthquake, whose moment magnitude is estimated at $M_w \sim 8.5$ (*Ambraseys and Douglas*, 2004; *Chen and Molnar*, 1977), could happen as often as once every 270 years within the borders of Nepal. As far as frequency is concerned, this would be the worst case scenario where all the moment deficit accumulated was released seismically in earthquakes whose magnitude wouldn't exceed 8.5. However, too many parameters remain unknown to make any accurate estimation on the return period of major earthquakes. Should the actual b-value of the seismicity distribution in Nepal slightly differ from 1, equation (1.8) shows that those estimates would be significantly affected. Another unknown parameter is the largest possible earthquake magnitude in Nepal which has a paramount effect, as shown on figure S6(b). Indeed if the seismicity on the MHT doesn't go beyond those $M_w \sim 8.5$ earthquakes, they would indeed have a period of return of about 270 years. But if we assume that the MHT can produce earthquakes up to $M_w \sim 9.2$, then the return period of $M_w \geq 8.5$ earthquakes would become of the order of 600 years.



(a) Recurrence time of the largest possible earthquakes in Nepal assuming a release of a proportion α of the accumulated moment by a seismicity following a Gutenberg-Richter distribution with $b = 1$.



(b) Reccurence time of $M_w \geq 8.5$ earthquakes as a function of the largest possible earthquakes happening in Nepal, for a moment accumulation of $\dot{\mathcal{M}}_0 = 6.6 \times 10^{19}$ Nm/yr released entirely seismically.

Figure S6: Estimations on the recurrence time of earthquakes.

Chapter 2

Detecting periodicities in earthquake catalogs using the Schuster test, application to Himalayan seismicity

Submitted to Earth and Planetary Science Letters, Jan. 2013

Thomas Ader,[1,2] Jean-Philippe Avouac[1]

[1]Department of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA 91125, USA.

[2]Laboratoire de Géologie, Ecole Normale Supérieure, CNRS, 24 rue Lhomond, 75004 Paris, France.

Abstract

We propose a method based on the Schuster test to search for periodicities in the timing of earthquakes in a catalog of independent events (i.e., a ‘declustered’ catalog). Such periodicities can be detected by computing a spectrum of Schuster p-values (the probability to observe such a level of periodic variations in a catalog occurring out of a constant seismicity rate), looking for harmonic variations of the seismicity rate at adequately chosen periods. We show that the detection level is actually period dependent, the 95% confidence detection level being achieved for Schuster p-values lower than $0.05 \times T/t$ rather than simply 0.05, where T is the period tested and t the duration of the catalog. Fortunately, this only translates into minor differences of amplitudes of seismicity rate variations for detection at the same confidence level. The Schuster spectrum is therefore an effective method to detect periodicities. It also provides information about the eventual non-harmonicity of the periodic signal in the catalog, or identifies an eventual imperfect declustering of the earthquake catalog, making it coincidently a potential tool to assess whether a catalog has been properly declustered. Applying this tool to the midcrudstal seismicity in Nepal, we show that intermediate magnitude events (i.e., $M_L \geq 5.5$ from the 1995 to 2008 NSC catalog and $M_b \geq 4$ from the 1965 to 2008 ISC catalog) exhibit annual variations of seismicity of amplitude of about 30%, while no other periodicity appears. In particular, no variations of seismicity at any of the tidal periods are observed. Seasonality of events at smaller magnitudes cannot be established with certainty by available data, as the seasonality of aftershocks subsequent to the seasonality of larger events may conceal or even replace it.

2.1 Introduction

On the road towards a global understanding of earthquake mechanics, numerous studies have added a paving stone by examining the response of seismicity to periodic stresses. The investigation is regularly carried out under the assistance of the Schuster test as a quantitative tool to either discard or reveal any existing correlation (*Schuster*, 1897; *Heaton*, 1975; *Tanaka et al.*, 2002a, 2006). Arthur Schuster first developed this test in 1897 (*Schuster*, 1897), using the derivations of *Rayleigh* (1880) to build a quantitative counter-argument to *Knott* (1897), who claimed that earthquakes and tides displayed a correlation in Japan. *Tanaka et al.* (2002a, 2006); *Cochran et al.* (2004) contemporarily revisited the question with modern catalogs, using the same test to shed light on some actual cases of tidal triggering.

The test considers the timing of events relative to the time variations of a perturbation, and ciphers out a p-value corresponding to the probability that the distribution of those relative times results from a uniform random process. It thereby provides a good measurement of the null hypothesis that events from a catalog do not correlate with a given periodic perturbation, and is therefore appropriate to investigate the correlation with any periodic forcings beyond tides. For instance, *Rydelek and Hass* (1994) used it to identify the presence of misidentified daily blasts in seismicity catalogs while *Bettinelli et al.* (2008) established the existence of annual variations of microseismicity in Nepal with it, which they linked to surface water load variations subsequent to the monsoon. *Lockner and Beeler* (1999) and *Beeler and Lockner* (2003) also used the Schuster test to quantify the response of a fault submitted to periodic load variations during lab experiments. All these studies thus used the Schuster test to determine how much the system responded to a known applied perturbation.

Conversely, we propose that the Schuster test may be used to identify periodicities of the seismicity rate in an earthquake time catalog. The idea is to compute a spectrum of Schuster p-values within a given range of periods, hereafter referred to as a Schuster spectrum, systematically tracking down hypothetical harmonic variations of the seismicity rate at the appropriate subset of periods.

After briefly presenting the principle of the Schuster test, section 2.3 determines the appropriate period sampling rate to build the Schuster spectrum as well as the threshold above which a peak in the spectrum can be regarded as significant. We then discuss the period dependence of the detection and artifacts that can occur in the spectrum in section 2.4, underlining the fact that the claim of periodic variations of the seismicity rate in a catalog requires the computation of a complete Schuster spectrum, rather than an isolated Schuster test. Finally, computing Schuster spectra for the midcrustal seismicity in Nepal in section 2.5, we show that annual variations of seismicity are discernible for the largest events of the catalogs, while they cannot be unequivocally claimed for small events. These annual variations are the only ones that come out of the spectra; in particular, no variations at the tidal periods are apparent.

2.2 The Schuster test

The Schuster test has been described in details in different studies (*Heaton*, 1975; *Rydelek and Hass*, 1994; *Tanaka et al.*, 2002a, 2006) and we here only summarize its pertaining principles in the case of harmonic variations. To compute the probability that the timing of events in a catalog varies harmonically at a period T, a phase is associated to each event: calling t_k the time of event number k, its associated phase θ_k is:

$$\theta_k = 2\pi \frac{t_k}{T}. \tag{2.1}$$

The catalog of times can hence be converted into a 2D walk made of successive unit length steps, in directions given by these phases. The probability p that a distance greater than or equal to D, the distance between the start and end points of this walk, can be reached by a uniformly random 2D walk is the probability of the null hypothesis that event times distribution arises from a uniform seismicity rate, and reads (e.g., *Schuster*, 1897):

$$p = e^{-D^2/N}, \tag{2.2}$$

where N is the number of events in the catalog. This probability is what we refer to as the Schuster p-value: the lower this p-value, the higher the probability of a periodicity at period T.

If a catalog contains N events occurring out of a harmonically varying seismicity rate:

$$\frac{R(t_k)}{r} = 1 + \alpha \cos\left(\frac{2\pi t_k}{T}\right), \tag{2.3}$$

where r is the average seismicity rate, α is the amplitude of the seismicity rate variations, t_k is the time of event number k and T the period of the variations, the logarithm of the

Schuster p-value computed at period T follows (see auxiliary material for derivation):

$$\begin{cases} \langle -\ln p \rangle = \left\langle \frac{D^2}{N} \right\rangle = 1 + \frac{N\alpha^2}{4}, \\ \mathrm{var}(\ln p) = \mathrm{var}\left(\frac{D^2}{N}\right) = \left(1 - \frac{\alpha^2}{2}\right)\left(1 + \frac{N\alpha^2}{2}\right). \end{cases} \tag{2.4}$$

The Schuster p-value is therefore independent of the period tested and only determined by both the number N of events in the catalog and the amplitude α of the seismicity rate variations.

2.3 Building a spectrum of Schuster p-values

The process of testing a periodicity in an earthquake catalog always boils down to the same underpinnings: the catalog gets stacked over the investigated period, and the probability that there exists a periodicity is subsequently estimated with whatever test is chosen (the Schuster test, fitting a sine-wave to the stacked catalog, etc.). Therefore, beyond the single period under investigation, the whole range of periodicities that remain coherent throughout the stacking process get actually tested.

Figure 2.1 illustrates this point. Let's suppose that we are searching for periodicities in a signal of length t that has a periodicity at period T (upper plot). We define the number of complete cycles in the catalog $n(T) = \mathcal{I}(t/T)$, where $\mathcal{I}(.)$ denotes the integer part of a real number. If the signal is stacked over the period T (lower left plot), then the periodicity is detected and the statistical test performed will quantitatively establish the existence of this periodicity. Now, let's imagine that instead of testing the period T, one tests a period $T + \Delta T_1$, such that $n\Delta T_1 \ll T$. In this case again, a periodicity appears in the stacked signal (lower middle plot), and is due to the periodicity at period T. The test at period $T + \Delta T_1$ is thus redundant with the one at period T since both will bring up the same periodicity. But if the period tested $T + \Delta T_2$ is such that the condition $n\Delta T_2 \ll T$ is not satisfied, then the periodic signal starts getting scrambled during the stacking process (lower right plot) and no periodicity will be detected.

Based on these considerations, one can determine the appropriate period sampling in order to be sure to test all periods within a given range: two consecutive tested periods T_i and $T_{i+1} = T_i + \Delta T_i$ have to verify $n(T_i)\Delta T_i < T_i$, or

$$n(T_i)\Delta T_i = \varepsilon T_i, \tag{2.5}$$

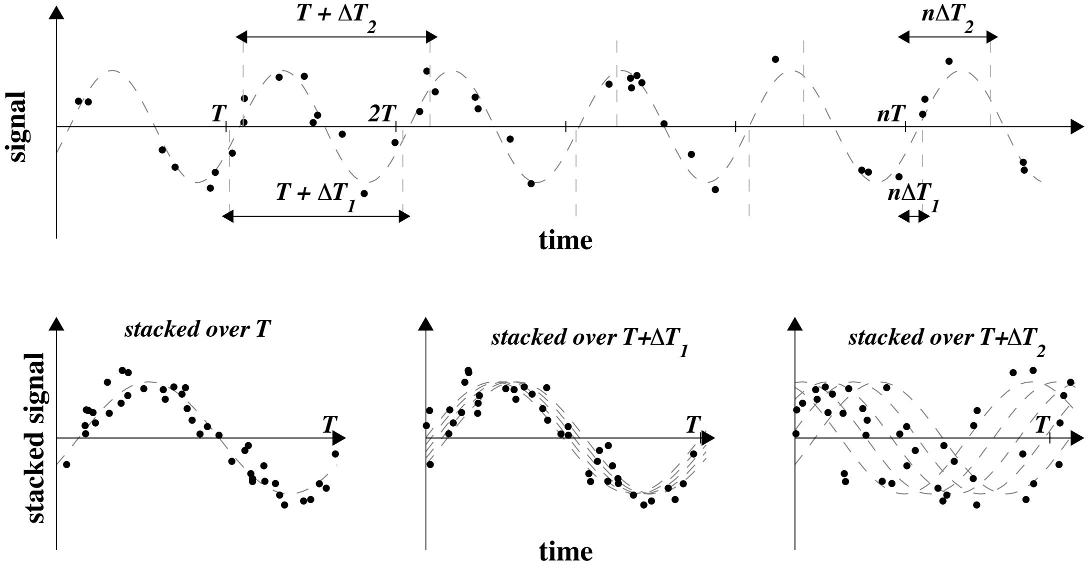


Figure 2.1: Figure showing why testing periodicities T and $T + \Delta T$ can be redundant or not, depending on the duration of the dataset. Upper plot: harmonically varying signal at period T over a duration t, such that $nT < t < (n+1)T$. Grey dashed line shows the signal while black dots represent hypothetical measurements. Errors have been added to hypothetical measurements, but those are not required for the reasoning to hold. Lower plots: from left to right, same signal stacked over different periods, respectively T, $T + \Delta T_1$ such that $n\Delta T_1 \ll T$, and $T + \Delta T_2$ such that condition $n\Delta T_2 \ll T$ is not fulfilled any more. In this last case, the periodicity at T is scrambled by the stacking process. It is therefore redundant to test periodicities both at T and $T + \Delta T_1$, but not at T and $T + \Delta T_2$.

where ε will be determined more precisely later. Noting that $n(T_i) = \mathcal{I}(t/T_i) \leq t/T_i$, the condition in equation (2.5) can be replaced by the following condition for the period increment:

$$\Delta T_i = \frac{\varepsilon T_i^2}{t}. \tag{2.6}$$

Noting $\nu = 1/T$ the frequency, the frequency increment is thus constant:

$$\Delta\nu = \frac{\varepsilon}{t}, \tag{2.7}$$

as would be the case for the set of frequencies at which a discrete Fourier transform would have to be evaluated for a classical time series with even spacing of data, in which case $\varepsilon = 1$ (e.g., *Scargle*, 1982; *Hernandez*, 1999).

Equation (2.6) shows that the period increment is smaller at short periods than at large periods. For instance, for a $t = 10$ year long earthquake catalog, taking $\varepsilon = 1$, in order to detect any periodicity around the main tides period (i.e., $T_{\text{tides}} \approx 0.5$ days), the period increment has to be $\Delta T_{\text{tides}} \approx 7 \times 10^{-5}$ days (≈ 6 s), while testing annual variations of seismicity only requires an increment $\Delta T_{\text{year}} \approx 37$ days.

Computing a spectrum between periods T_{min} and T_{max} (i.e., between frequencies $\nu_{\text{min}} = 1/T_{\text{max}}$ and $\nu_{\text{max}} = 1/T_{\text{min}}$) requires performing N Schuster tests, where from equation (2.7):

$$N = \frac{t}{\varepsilon}\left(\frac{1}{T_{\text{min}}} - \frac{1}{T_{\text{max}}}\right) \approx \frac{t}{\varepsilon T_{\text{min}}}, \tag{2.8}$$

since in general $T_{\text{min}} \ll T_{\text{max}}$.

As is suggested by equation (2.6), the subset of periods at which the spectrum will be computed depends on the choice of ε. Choosing a value too large, the spectrum will "miss" some periods, while values too small will yield an oversampling of periods and consequently an unnecessarily long computation time. We thus determine ε_o, the optimal value of ε, which is the largest value of ε such that all periods are tested, for a subset of periods built according to equation (2.6). Deriving a Schuster spectrum for a collection of periods between T_{min} and T_{max} with $\varepsilon > \varepsilon_o$, the N periods tested will hence be independent. As a result, the probability that all N Schuster p-values computed in the spectrum are greater than a given value δ is:

$$P = (1 - \delta)^N. \tag{2.9}$$

Since in general $N \gg 1$, for $\delta \ll 1$ (otherwise $P \approx 0$), equation (2.9) can be well approximated by:

$$P = e^{-\delta N}. \tag{2.10}$$

Calling δ_m the smallest Schuster p-value of the spectrum, the random variable $X = N\delta_m$ has thus a Poissonian probability density function (hereafter PDF):

$$p_X(X) = e^{-X}, \tag{2.11}$$

and an expected value $\langle X \rangle = 1$, simply reflecting the fact that for a random catalog (with no periodicity of the seismicity rate), the Schuster p-value has a uniform PDF over [0;1].

In order to estimate the actual value of ε_o, we build the PDF of the random variable

$$Y = \varepsilon X = \delta_m \left(\frac{t}{T_{\text{min}}} - \frac{t}{T_{\text{max}}}\right) \tag{2.12}$$

which has the following PDF if $\varepsilon > \varepsilon_o$:

$$p_Y(Y) = \frac{1}{\varepsilon}\, e^{-Y/\varepsilon}, \tag{2.13}$$

since in this case all periods tested in the spectrum are independent. In the case $\varepsilon < \varepsilon_o$, the periods tested are not independent any more, and the actual number of independent periods tested becomes $N_c \leq N$ since some periods are redundant. The PDF of the random variable Y thus becomes:

$$p_Y(Y) = \frac{1}{\varepsilon_c}\, e^{-Y/\varepsilon_c}, \tag{2.14}$$

where $\varepsilon_c \geq \varepsilon$ is such that $N_c = N(\varepsilon_c)$, following equation (2.8).

In order to estimate the optimal value ε_o, we build the PDF of the random variable Y for different values of ε, by computing Schuster spectra on randomly generated catalogs, and fit each PDF with a law following equation (2.14), fitting for the parameter ε_c. If the initial choice of ε is smaller than ε_o, the periods tested in the spectrum are not independent, and the best fit will thus be obtained for a value of ε_c greater than ε. On the other hand, if the initial value of ε is large enough for all periods tested to be independent, the best fit will be obtained for $\varepsilon_c = \varepsilon$. The optimal value ε_o is therefore the minimum value of ε_c such that $\varepsilon_c = \varepsilon$.

Figure 2.2 illustrates this process. Figure 2.2a shows a plot of the PDF of Y computed with 10^5 randomly generated catalogs of 10^4 events each, for each of which the Schuster spectrum has been evaluated from $T_{\min}/t = 10^{-3}$ to $T_{\max}/t = 0.1$ and $\varepsilon = 0.1$. We derive the complete PDF of ε_c (inset in Figure 2.2) by computing the probability that equation (2.14) represents the PDF of Y for different values of ε_c. In the case of Figure 2.2a, this leads to $\varepsilon_c = 0.357 \pm 0.002$. Technical details on this process are given in the supplementary material, section S.2.

This operation is then repeated for different initial values of ε and the obtained values of ε_c are plotted in Figure 2.2b, as a function of the initial value ε. Figure 2.2a thus corresponds to the leftmost point on Figure 2.2b. Figure 2.2b shows that the periods tested seem to start being independent for $\varepsilon \geq 0.8$. Hereafter, we simply choose to use $\varepsilon_o = 1$.

With the value of ε_o in hand, and thus the PDF for the minimum Schuster p-value, it is possible to estimate the expected threshold above which a Schuster p-value will indicate with confidence that the seismicity rate contains a periodicity. From equation (2.11), the expected value of the minimum Schuster probability is $\langle \delta_m \rangle = 1/N \approx \varepsilon_o T_{\min}/t \approx T_{\min}/t$. Since throughout the spectrum, in general $T \ll T_{\max}$, the expected value of the minimum

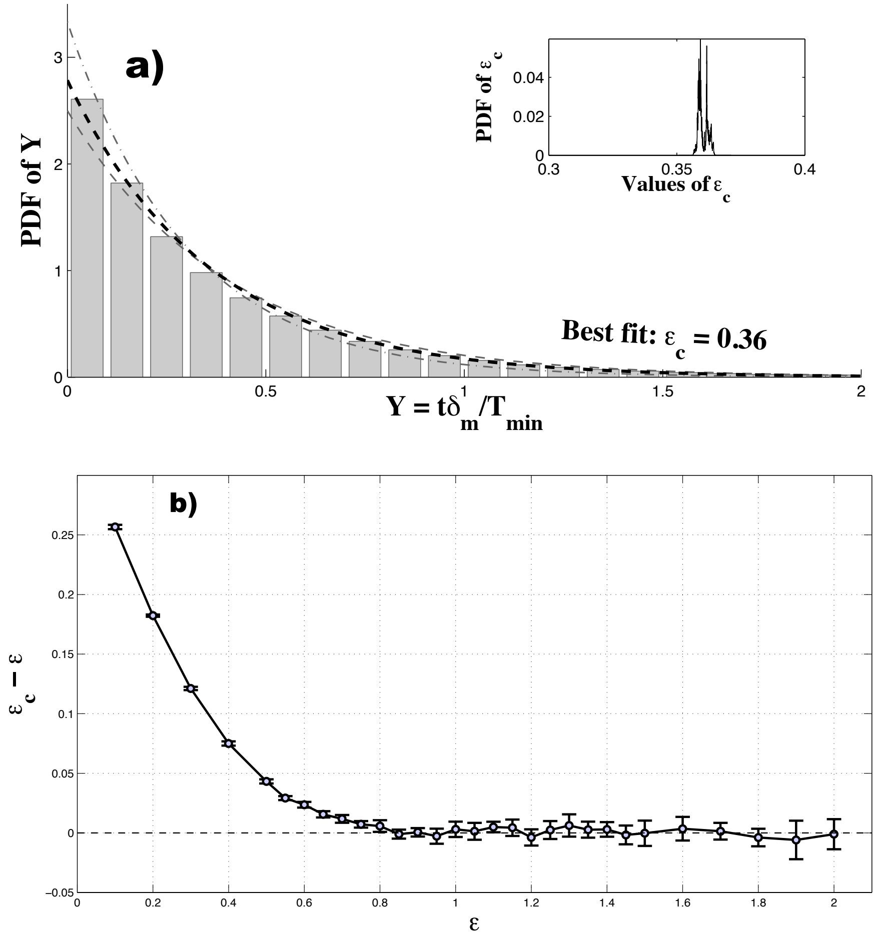


Figure 2.2: a) PDF of random variable $Y = t\delta_m/T_{\mathrm{min}}$ (bar plot) and fit with equation (2.14) for different values of ε_c. The PDF was simulated with 10^5 randomly generated 10^4 event catalogs, taking $\varepsilon = 0.1$, and computing the Schuster spectra between $T_{\mathrm{min}}/t = 10^{-3}$ and $T_{\mathrm{max}}/t = 10^{-1}$, where t is the total length of the catalog. Dashed black line corresponds to the best fitting model ($\varepsilon_c = 0.36$), while grey lines correspond to the fit for $\varepsilon_c = 0.3$ (dash-dotted line) and $\varepsilon_c = 0.4$ (dashed line). Inset: complete PDF of ε_c. b) Difference between ε_c (obtained from the fit to the PDF as is shown on figure a) and input value of ε to compute the spectrum, for different initial value of ε. Error bars represent the 1-σ uncertainties obtained from the complete PDF of ε_c.

Schuster probability for periods greater than T is simply:

$$\langle \delta_m \rangle = \frac{T}{t}. \tag{2.15}$$

A periodicity in the catalog will thus have a significant probability to exist if its Schuster p-value is significantly lower than this expected value. Quantitatively, a periodicity can be

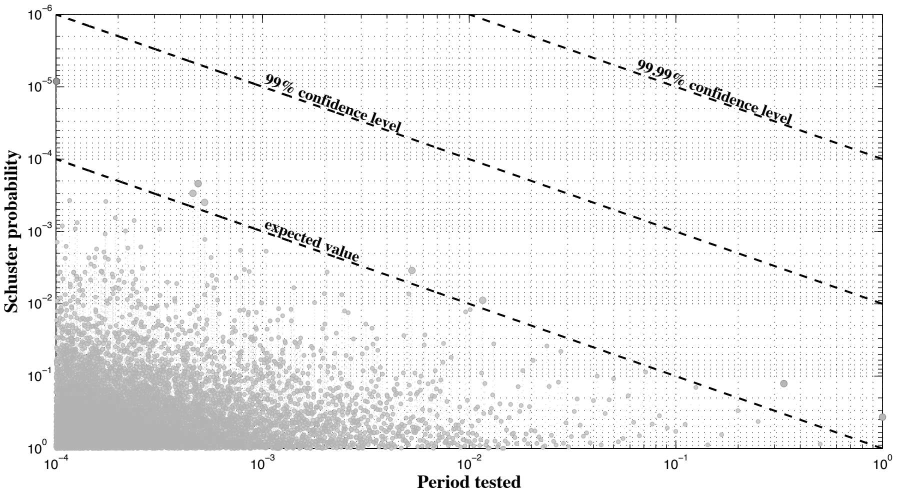


Figure 2.3: Schuster spectrum over a uniformly random 1000 event catalog of length t, computed between $T_{\min}/t = 10^{-4}$ and $T_{\max} = t$. The period dependent "expected value" dashed line represents the expected minimum Schuster p-value, and corresponds to equation (2.15), while "99% confidence level" corresponds to 1% of minimum expected values.

claimed to be detected above the 95% confidence level if the corresponding Schuster p-value is lower than $0.05 \times \langle \delta_m \rangle = 0.05 \times T/t$, rather than simply 0.05: the detection level is thus period dependent, being better at larger periods. Figure 2.3 shows the Schuster spectrum obtained for a 1000 event catalog of duration t, generated out of a uniform seismicity rate. The spectrum is built between $T_{\min}/t = 10^{-4}$ and $T_{\max} = t$. Even though the catalog does not contain any periodicity, the Schuster test returns smaller Schuster p-values at short periods (equation (2.15)), due to a greater density of periods tested.

As suggested by equation (2.7), a "flat" spectrum would be obtained for a linear x-axis in frequencies. However, with such a representation, the expected value of the Schuster p-values could not be represented by a simple straight line anymore.

A periodicity in the seismicity rate thus requires a lower Schuster p-value for the detection to be considered significant a shorter periods. However, if equation (2.15) may suggest a drastic dependence of the detection threshold on the period, combining it with equation (2.4) leads to the following expression for the critical amplitude of seismicity rate variations necessary for a detection at the 95% confidence level:

$$\alpha_{95} = \frac{2}{\sqrt{N}} \sqrt{2 + \ln \frac{t}{T}}, \tag{2.16}$$

indicating that the critical amplitude of the seismicity rate variations above which a periodicity can generally be detected is not very sensitive to the period. For a 1000 event catalog covering 10 years, $\alpha_{95}(T = 1 \text{ year}) \approx 14\%$, while $\alpha_{95}(T = T_{\text{tides}}) \approx 21\%$. The difference thus remains minor, suggesting that the Schuster spectrum offers a reliable way to detect periodicities.

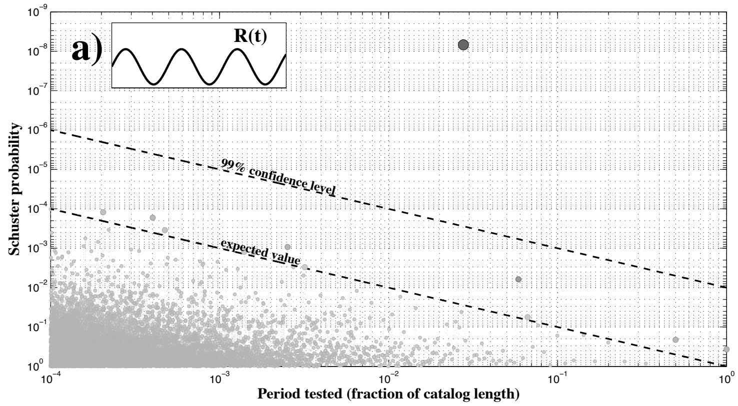
2.4 Application to synthetic catalogs

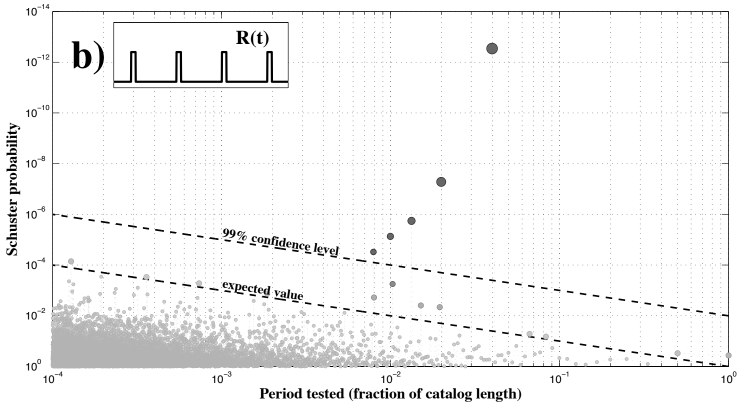
Now that we have exposed how to build a Schuster spectrum and established the levels of confidence for detection of periodicities, we apply it to three different synthetic catalogs, in order to show that it is able to detect an unknown periodicity, and that it actually is the only way to claim whether the catalog analyzed contains a periodicity or not. Indeed, if a catalog occurs out of a uniform seismicity rate, or if the variations of seismicity rate are randomly distributed with respect to the period tested, the Schuster test will return high p-values. However, if variations of the seismicity rate remain coherent throughout the stacking process, the Schuster p-values will be small, whether the variations of seismicity rate are periodic at the period considered or not. In particular, periodic variations of the seismicity rate at periods that are an integer multiple of the period considered, or a sudden outburst of seismicity will lead to low p-values. We thus apply the Schuster spectrum to the three following types of catalogs: one generated out of a harmonically varying seismicity rate, one out of a periodic but non-harmonic seismicity rate, and one out of a uniform seismicity rate superimposed with an aftershock sequence. Supplementary material section S.3 describes in details how the catalogs have been generated.

First of all, Figure 2.4a represents the Schuster spectrum from a 1000 event catalog generated from a harmonic seismicity rate following equation (2.3), with $T/t = 0.029$ and $\alpha = 0.35$. In this case, the spectrum clearly reveals the periodicity at period T/t, and no other periodicity appears.

However, if the seismicity rate is periodic, but non harmonic, harmonics of the main periods may appear in the spectrum. Figure 2.4b shows the spectrum for a catalog generated out the following periodic, non-harmonic, seismicity rate:

$$\frac{R(t_k)}{r} = \begin{cases} 1, & \text{if} \quad t_k[T]/T \in [0; 0.1] \\ \alpha, & \text{if} \quad t_k[T]/T \in [0.1; 0.2] \\ 1, & \text{if} \quad t_k[T]/T \in [0.2; 1] \end{cases}, \tag{2.17}$$


a)
R(t)
Schuster probability
Period tested (fraction of catalog length)
99% confidence level
expected value
10^{-9}
10^{-8}
10^{-7}
10^{-6}
10^{-5}
10^{-4}
10^{-3}
10^{-2}
10^{-1}
10^{0}
10^{-4}
10^{-3}
10^{-2}
10^{-1}
10^{0}


b)
R(t)
Schuster probability
Period tested (fraction of catalog length)
99% confidence level
expected value
10^{-14}
10^{-12}
10^{-10}
10^{-8}
10^{-6}
10^{-4}
10^{-2}
10^{0}
10^{-4}
10^{-3}
10^{-2}
10^{-1}
10^{0}

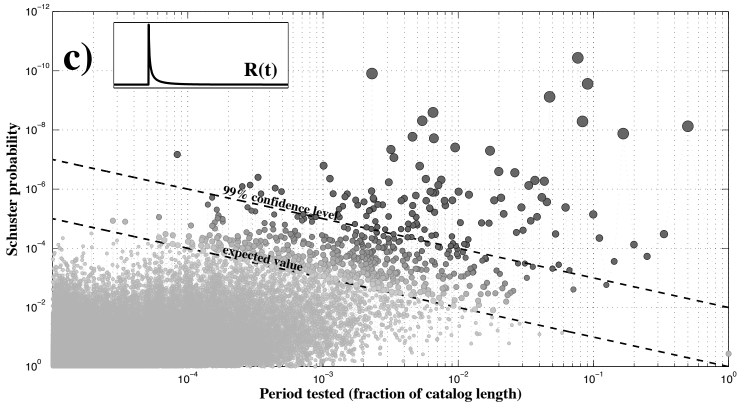

c)
R(t)
Schuster probability
Period tested (fraction of catalog length)
99% confidence level
expected value
10^{-12}
10^{-10}
10^{-8}
10^{-6}
10^{-4}
10^{-2}
10^{0}
10^{-4}
10^{-3}
10^{-2}
10^{-1}
10^{0}

Figure 2.4: a) Schuster spectrum for a 1000 event catalog generated out of a harmonically varying seismicity rate following equation (2.3), with variations of amplitude $\alpha = 0.35$, and period $T/t = 0.028$, where t is the total duration of the catalog. b) Schuster spectrum for a 1000 event catalog generated out of a periodic, non-harmonic seismicity rate following equation (2.17), with variations of amplitude $\alpha = 3$, and period $T/t = 0.04$. c) Schuster spectrum for a 1000 event catalog containing a *Dieterich* (1994) type aftershock sequence, which seismicity rate is given by equation (2.18), with characteristic aftershock decay time $t_a/t = 10^{-2}$. Other parameters are given in the main text. The inset at the top left corner of each spectrum schematically represents the seismicity rate used to generate the catalog.

where t_k is the time of event number k, T is the period of the seismicity rate, $t_k[T]$ is the modulus of t_k after division by the period T, and $\alpha > 1$ a manually chosen parameter. A schematic of this seismicity rate is shown as an inset in Figure 2.4b. The seismicity rate used to generate the catalog analyzed in Figure 2.4b was obtained for $T/t = 0.04$ and $\alpha = 3$. In this case, if the periodicity at $T/t = 0.04$ appears clearly, one also notices harmonic periodicities in the spectrum at $T_2/t = 0.02$, $T_3/t = 0.0133$ and $T_4/t = 0.01$. It is easy to understand how such periodicities show up in the spectrum from the way the Schuster test works: when stacked over the period T, the seismicity rate is $\alpha = 3$ times higher at times between $0.1 \times T$ and $0.2 \times T$. When stacked over the period $T/2$, it is 3 times higher every other cycle for times between $0.2 \times T/2$ and $0.4 \times T/2$, and is thus on average 2 times higher on that interval of times. When stacked over the period $T/3$, it is 3 times higher every 3 cycle for times between $0.3 \times T/3$ and $0.6 \times T/3$, and is thus on average 1.667 times higher on that interval of times. This reasoning can be applied to all successive harmonics, until the time span over which the seismicity rate is higher becomes of the order of the harmonic's period. Those harmonics should thus be disregarded when looking for independent periodicities in the catalog, but the smallest harmonic appearing in the spectrum provides an estimate of the duration of the higher seismicity rate within one period.

Another configuration of catalog that might lead to a bad period detection is the case where some events are not independent from each other and cluster in time, as is the case for instance if the catalog contains an aftershock sequence. In this case, keeping these aftershocks in the catalog might conceal some periodic variations in the background seismicity rate. This is illustrated by the spectrum on Figure 2.4c: the simulated catalog has a uniform background seismicity rate r with an aftershock sequence superimposed to it, and its seismicity rate is described by the following law (*Dieterich*, 1994):

$$\frac{R(t_k)}{r} = \frac{1}{1 + (e^{-Q} - 1)\, e^{-(t_k - t_m)/t_a} \mathcal{H}(t_k - t_m)}, \tag{2.18}$$

where t_k is the time of event number k, the background seismicity rate r is supposed to be identical before and after the aftershock sequence, e^Q is the normalized seismicity rate right after the mainshock, t_m is the time of the mainshock and t_a the characteristic duration of the aftershock sequence. The function $\mathcal{H}(.)$ is the Heavyside function ($\mathcal{H}(x) = 0$ for $x < 0$ and $\mathcal{H}(x) = 1$ for $x \geq 0$). In order to make sure that the aftershock sequence is over before the end of the catalog, we suppose that $t_a \ll t - t_m$. The catalog used to generate Figure 2.4c contains a background of 1000 events and an aftershock sequence containing 100 events, with $t_a = 10^{-2}t$, $t_m = 0.2t$ and $Q = 10$. In this case, for all periods of the order of or larger than the characteristic duration of the aftershock cluster, the "Schuster walk" will progress in one direction by a large distance during the aftershock sequence, systematically resulting in artificially low Schuster p-values that might conceal existing periodicities of the background rate. The same thing will happen for a swarm of earthquakes in a catalog or any increase of seismicity rate over a duration less than the period tested. This misinterpretation of clusters into periodic variations is not inherent to the Schuster test itself, it only comes from the stacking of events times over the period considered. Deriving the entire spectrum thus provides a mean to detect if a low Schuster p-value might be due to clusters. Conversely, the Schuster spectrum might also be used to assert if the catalog contains clusters, whatever their nature, in which case the spectrum will systematically display low p-values at large periods.

These tests on synthetic catalogs show that the Schuster spectrum proves to be an efficient tool to detect unknown periodicities in the seismicity rate of an event catalog, but also outline the paramount benefit of the whole Schuster spectrum over a isolated Schuster test. It provides a much more precise diagnostic on whether a catalog contains a periodicity, or if low Schuster p-values are due to different non-uniformities of the seismicity rate.

2.5 Application to the seismicity of Nepal

The seismicity of the Nepal Himalaya has been reported to undergo seasonal variations of its rate (*Bollinger et al.*, 2007; *Bettinelli et al.*, 2008), and is a good case of study to apply the Schuster spectrum.

A large fraction of the earthquakes in Nepal cluster at the downdip end of the locked part of the Main Himalayan Thrust (MHT) fault (*Cattin and Avouac*, 2000; *Bollinger et al.*, 2004; *Ader et al.*, 2012a), forming a belt of seismicity at the front of the Himalayan chain (*Pandey et al.*, 1995), well recorded by the National Seismological Center (NSC) in Kathmandu, Nepal. Looking at events from this midcrustal cluster from 1995 to 2000, *Bollinger et al.* (2007) reported seismicity rates 30% to 60% higher during the winter than the summer months, which they attributed to stress variations on the MHT subsequent to surface load variations following the hydrological cycle.

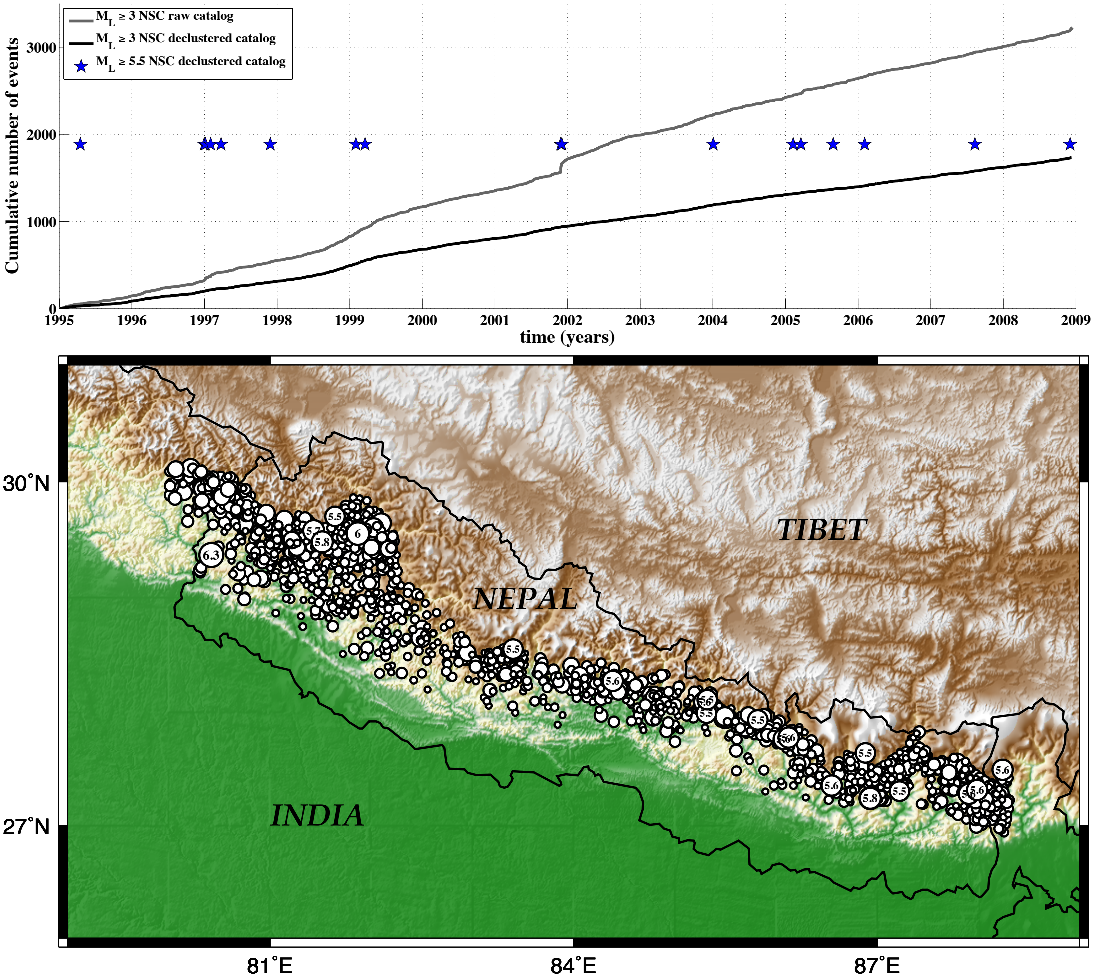


Figure 2.5: Time and space distribution of the $M_L \geq 3$ NSC seismicity used in this study. Upper plot shows cumulative number of events from 1995 to the end of 2008 for raw (grey curve) and declustered (black curve) catalogs, together with times of $M_L \geq 5.5$ events from the declustered catalog (blue stars). The map shows midcrustal events from the raw catalog used in the study, selected according to their localization, using the same selection contour as in *Bollinger et al.* (2007). Circles sizes are proportional to events magnitudes: smallest events have $M_L = 3$, and $M_L \geq 5.5$ events are indicated by their magnitude, giving an idea of the scale.

Today, the available seismicity catalog compiled by the NSC extends until the end of 2008. This, together with the Schuster spectrum presented above, provides new material to both reassess the significance of the reported seasonal variations of seismicity in Nepal, and look for any other periodic variations of the seismicity rate. In parallel, we corroborate our results by independently looking at the ISC catalog (*International Seismological Centre*, 2010) from 1965 to 2008. For both catalogs, we isolate events from the midcrustal cluster using the same contour in map view as in *Bollinger et al.* (2007), and decluster with the algorithm described in *Reasenberg* (1985), with the same set of parameters as in *Bollinger et al.* (2007) ($P = 0.95$, $1 \leq \tau \leq 10$ days, $D \leq 20$ km, $U_x = 5$ km and $U_z = 10$ km). Figure 2.5 shows the temporal evolution of $M_L \geq 3$ selected events from both the raw and the declustered NSC catalogs, together with a map showing their spatial distribution (the map showing the position of ISC events is available in the supplementary material, Figure S2).

Burtin et al. (2008) showed that the seismic noise at the recording seismic stations, largely imputable to friction of pebbles at the bottom of rivers, was higher in the summer, due to higher water stream power and discharge. In order to avoid any contamination of our results by these seasonal variations of seismic noise, we consider only events with local magnitude $M_L \geq 3$, which is above the detection level at all time (*Bollinger et al.*, 2007). We first derive the Schuster spectrum for the catalog used in the study by *Bollinger et al.* (2007), both before and after declustering (respectively figures 2.6a and 2.6b). The Schuster spectrum of the catalog before declustering (Figure 2.6a) displays a prominent peak at one year, but it also contains numerous peaks at larger and smaller periods, indicating, as has been showed in the previous section with synthetic catalogs, that the observed annual periodicity may as well be due to clusters present in the catalog, such as aftershock sequences or other abrupt changes of seismicity rate (Figure 2.5). There is, for instance, a global increase of seismicity rate by a factor of more than 2 during the winter months of 1998-1999 (see also Figure S3, supplementary material), even after declustering, which would induce low p-values at periods larger than about a year. Once the catalog is declustered (Figure 2.6b), most of the previous periodicities disappear from the spectrum. The annual period returns a Schuster p-value between the 95% and 99% confidence levels, but can hardly be claimed as a clear annual periodicity of the seismicity rate rather than the product of clusters in the catalog, since periodicities at larger periods consistently remain. Given the 6 year total duration of the catalog, it is difficult to assert whether these periodicities are valid or also due to clusters in the catalog, such as the one during the winter months of

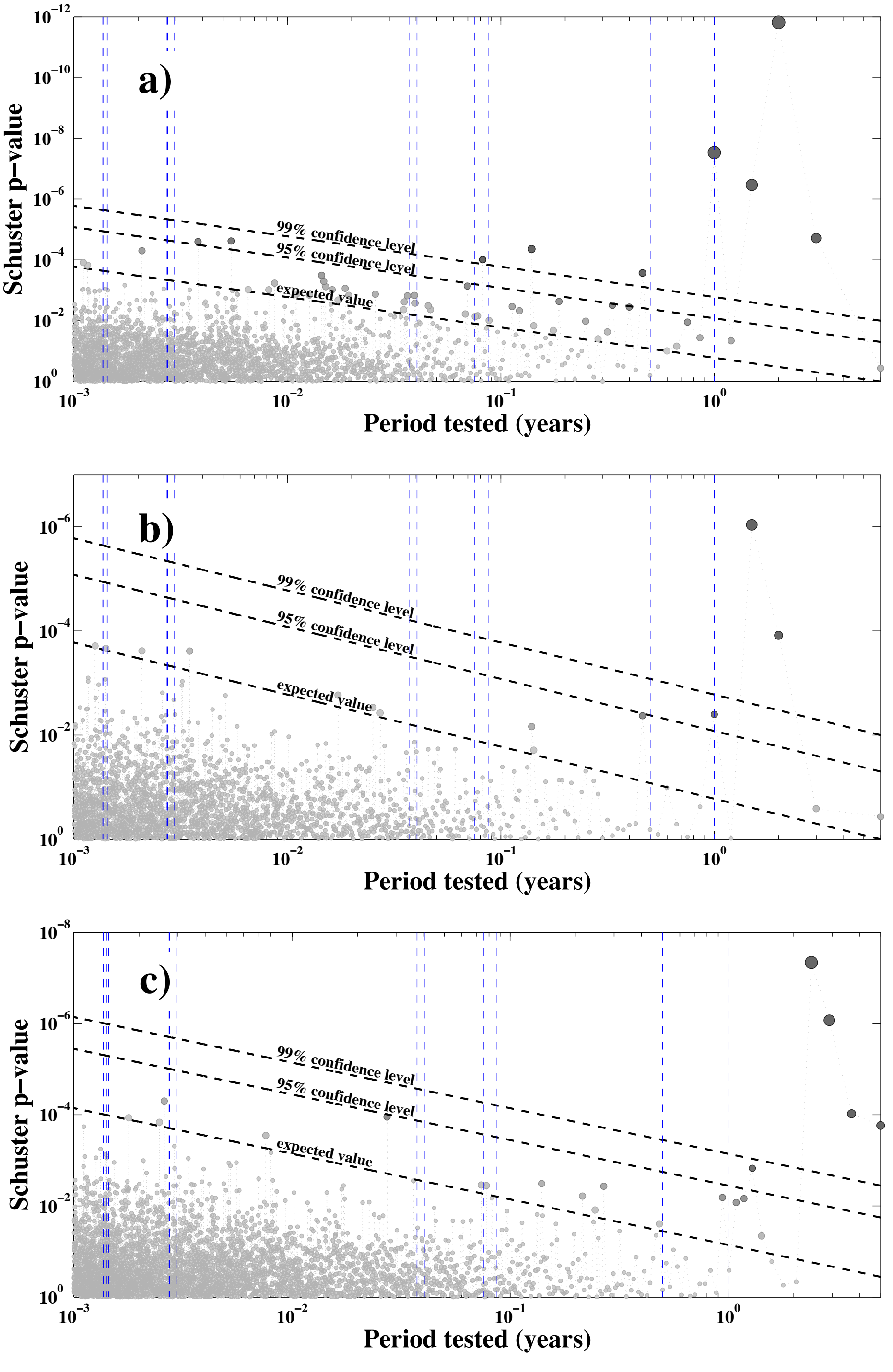


Figure 2.6: Schuster spectra built with the NSC $M_L \geq 3$ seismicity catalog over the same period of time as in *Bollinger et al.* (2007) (i.e., from 1995 to 2001), a) before and b) after declustering, and c) for the whole available NSC $M_L \geq 3$ declustered catalog (i.e., from 1995 to 2008). The vertical dashed blue lines indicate tidal, half annual and annual periods.

1998-1999. Declustered or not, this catalog does not display any variation of the seismicity rate at any of the tidal periods.

Extending the $M_L \geq 3$ NSC catalog until the end of 2008, the peak at 1 year actually disappears from the spectrum (Figure 2.6c), while peaks at periods larger than 2.5 years consistently remain, suggesting that clusters are most likely still present. This outlines the main drawback of working with small magnitude events: although they come in a statistically significant number, they easily violate the independence hypothesis, owing to their sensitivity to local perturbations, which could originate from larger seismic events or other possible forcing events (slow slip events, sub-surface hydrology, mining, etc.). Moreover, if both large and small events follow similar periodic variations of seismicity rate, small events' rate variations might be occulted by larger events' aftershocks, and thereupon ironed out during the declustering process.

Deriving the whole Schuster spectrum on these catalogs thus shows that no variations at periods less than a year are manifest, but that no rigorous claim can be made for periods of 1 year or more, because of clusters present in the catalogs.

A natural way to circumvent this issue is therefore to look at events of larger magnitude. This is now possible thanks to the longer time span of the NSC declustered catalog, which contains 16 events of $M_L \geq 5.5$ from 1995 to 2008 (up to $M_L = 6.3$), only 3 of which happen around the summer months (decimal year between 0.25 and 0.75, see blue stars on Figure 2.5), a misbalance that only has a 2% binomial probability of happening out of a uniform seismicity rate. On the Schuster spectrum computed for these events, the annual periodicity appears as the only one above the 95% confidence level (Figure 2.7a), suggesting indeed that peaks at other periods on the spectrum for $M_L \geq 3$ events were mostly due to singular variations of the seismicity rate. Besides, here again, no variations at the tidal period stand out.

The Schuster spectrum computed over the 210 $M_b \geq 4$ events from the declustered ISC catalog taken from 1965 to 2008 (Figure 2.7b) backs up these observations: the periodicity at 1 year still emerges alone above the 95% detection level. Equation (2.4) indicates that the Schuster p-value at 1 year corresponds to variations of the seismicity rate of $\alpha = 27\% \pm 7\%$, a value close to the one claimed in *Bollinger et al.* (2007). The $M_b = 4$ magnitude selection threshold may seem low, especially in the earlier years of the catalog, but since the completeness magnitude does not vary in a periodic way for the ISC catalog, this would not affect the detection of periodicities. Annual variations of seismicity thus prevail for larger events, and stand alone as the only periodic variations of the midcrustal seismicity

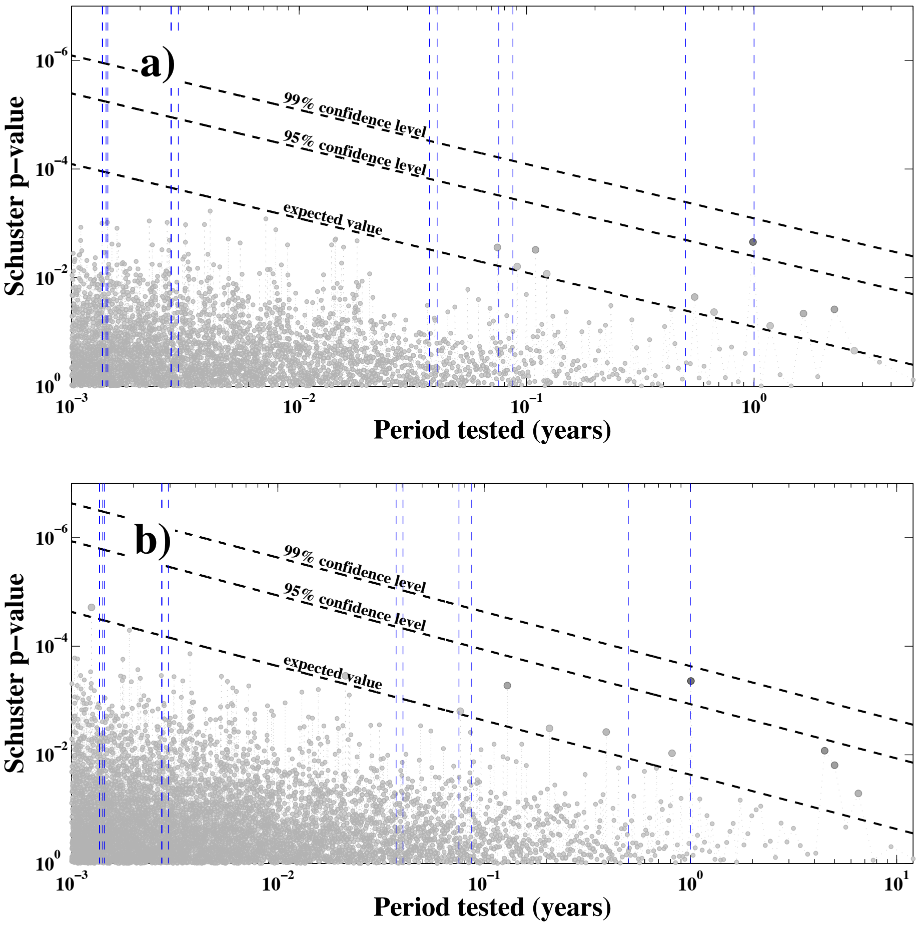


Figure 2.7: Schuster spectrum computed for large events from a) the NSC and b) the ISC declustered catalogs. For the NSC catalog, the spectrum is computed for $M_L \geq 5.5$ events for the entire available catalog (1995 to 2008). The spectrum for the ISC catalog is computed for $M_b \geq 4$ events from 1965 to 2008.

in Nepal. In particular, no variations are detected at any of the tidal periods.

Using equations (2.4) and (2.15) together with current seismicity rates from the NSC catalog, Figure 2.8 shows the minimum theoretical catalog duration necessary to detect annual variations of seismicity at the 95% confidence level for the NSC catalog at different cutoff M_L. It indicates that 6 years of the declustered $M_L \geq 3$ NSC catalog should be enough to detect variations of the seismicity rate of amplitude greater than 20%, and that the 14 available years of this catalog should actually enable us to detect variations of amplitude as low as 15% (or 22% when looking at $M_L \geq 3.5$ events). Were the annual variations of the seismicity rate of the $M_L \geq 3$ NSC events as intense as those of the $M_b \geq 4$ ISC events (i.e., 27% ± 7%), they should thus clearly come out of the spectra in Figure 2.6. However,

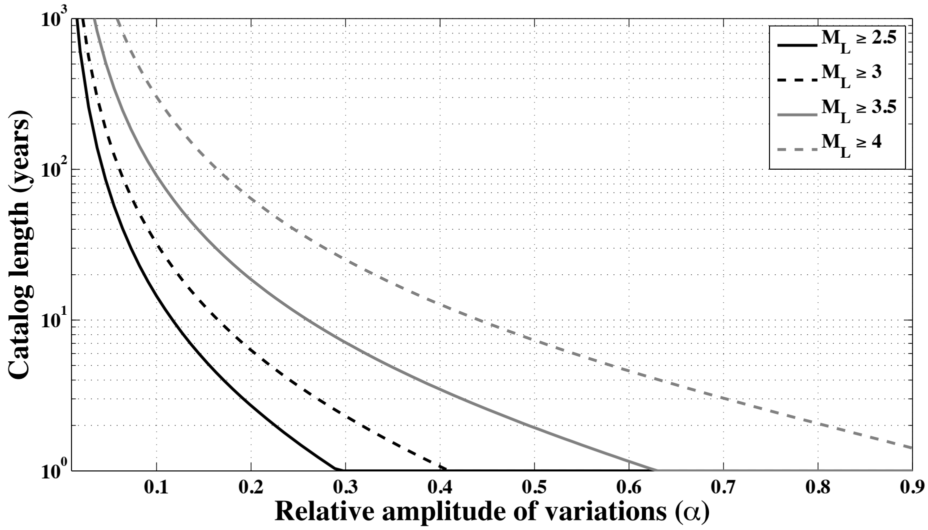


Figure 2.8: Theoretical minimum duration of the NSC catalog in order to detect variations of seismicity rate of a given amplitude α at the 95% confidence level, for different cutoff magnitudes. This theoretical duration is solution of the equation $\ln p + \sigma_{\ln p} = \ln(0.05 \times \delta_m)$, where $\ln p$ and $\sigma_{\ln p}$ are given by equation (2.4), and δ_m is given in equation (2.15).

once the NSC catalog is declustered, they might appear when considering the catalog up to 2001 (6 years of data), although they cannot be told apart from the presence of clusters of seismicity, and clearly do not show up anymore in the 1995-2008 catalog.

As has been discussed before, one possible reason is that the seasonal variations of seismicity for small events get smoothed out during the declustering process. This would explain in retrospect why annual variations of seismicity appear so clearly in the undeclustered catalog (Figure 2.6a), although once again this is rigorously not possible to assert it with this spectrum. Large winter events may in fact trigger surrounding faults close to failure, that would have otherwise ruptured later in the winter, thus annihilating the gradual increase of seismicity. In other words, these small events occur all at once as aftershocks of a larger event instead of as the result of a slow increase of seismicity.

Another possibility for this observation, is that the amplitude of the seasonal variations might vary with time, as has been proposed for earthquakes triggered by earth tides (*Tanaka et al.*, 2002b). Figure 2.9 shows the evolution of the Schuster p-value at 1 year when successively adding years to the $M_b \geq 4$ declustered ISC catalog, as well as the theoretical Schuster p-value with the 1σ standard deviation (equation (2.4)), for variations of amplitude $\alpha = 40\%$, given the number of events in the catalog. Note that the value $\alpha = 27\% \pm 7\%$ specifically corresponds to the Schuster p-value considering the whole 1965-2008 period

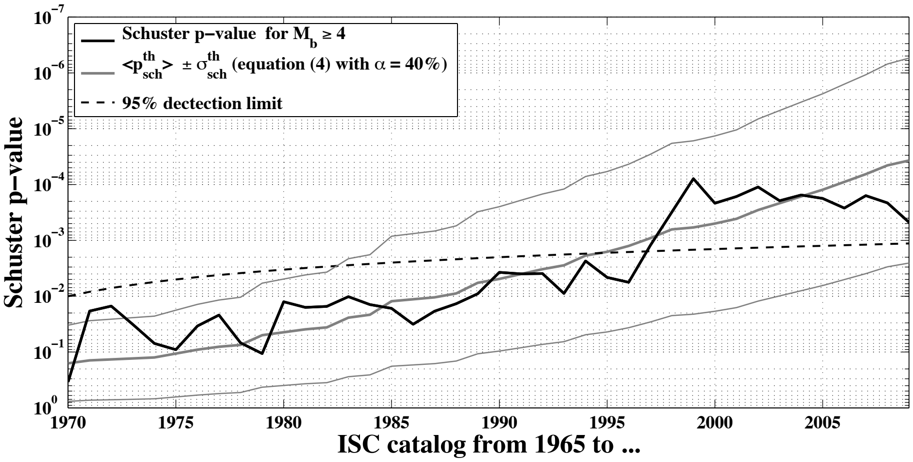


Figure 2.9: Evolution of the Schuster p-value by adding years to the ISC declustered catalog, considering $M_b \geq 4$ events, starting in 1965. The dashed black line indicates the detection at the 95% confidence level. The grey lines show the expected Schuster p-value for the ISC catalog given the number of events in the catalog, assuming annual variations of the seismicity rate of amplitude $\alpha = 40\%$, using equation (2.4).

(rightmost value of the black curve in Figure 2.9), while the value $\alpha = 40\%$ seems to better follow the general trend of evolution of the annual Schuster p-value in Figure 2.9. In this plot, the magnitude of the annual variations of the seismicity rate seems to decrease after 2000, but the p-value remains within one standard deviation of its expected value. This indicates that the decrease might simply be a statistical effect: the amplitude of annual variations of seismicity rate may globally be around 40%, but may appear to be less when looking at a short time range. This decrease seems to also appear in the $M_L \geq 3$ declustered NSC catalog (supplementary figure S4), although Figure S4 can be misleading since as has been explained earlier, the Schuster p-value at 1 year is also affected by isolated variations of the seismicity rate at such low magnitudes.

2.6 Conclusion

We propose a way to use the Schuster tests in order to build an entire spectrum of Schuster p-values, testing for harmonic variations of seismicity at a properly selected set of periods. The obtained spectrum provides an efficient tool to both detect unknown periodicities in an earthquake catalog and assert if variations of seismicity rate in the catalog are actually periodic or not, something that an isolated Schuster test cannot do.

Applying this Schuster spectrum to earthquakes catalogs from the midcrustal cluster of seismicity in Nepal suggests that intermediate events ($M_L \geq 5.5$ or $M_b \geq 4$) exhibit seasonal variations of seismicity, with an increase of seismicity in the winter of about 30%. The complete spectrum shows that seasonal variations of seismicity at lower magnitudes cannot be established with the same confidence, as aftershock sequences are more numerous in the winter and might therefore cover an increase of background seismicity. Whatever magnitude one examines though, no other periodic variations of seismicity rate appear in the catalog. In particular, no periodicity at any of the tidal periods is detected.

The implementation of the spectrum is straightforward. We propose an implementation written in Matlab, which can be found on the Tectonics Observatory's website (`http://www.tectonics.caltech.edu/resources`). The code `Schuster_test_log.m` computes the log p-value for a given catalog at a given array of periods, while the code `Schuster_spectrum.m` computes and plot the whole Schuster spectrum of a catalog between 2 given periods.

Acknowledgments

We thank Francisco Ortega for his help in writing some of the algorithms for the simulations. This project was supported by the Gordon and Betty Moore Foundation, through the Tectonics Observatory, NSF grant EAR #0838495. This is Caltech Tectonic Observatory's contribution number XXX and Caltech Seismolab contribution number XXX.

Supplementary material

S.1 Mean and variance of the Schuster ln$p$ value

It is possible to analytically tie together the Schuster p-value and the parameters of a normalized seismicity rate with harmonic variations:

$$\frac{R(t_k)}{r} = 1 + \alpha \cos\left(\frac{2\pi t_k}{T}\right), \tag{S1}$$

where α is the amplitude of the seismicity rate variations, t_k is the time of event number k and T the period of the variations. The Schuster test corresponds to N steps in directions θ_k defined in equation (2.1), where the PDF of each of the random variable θ_k is, from equation (S1)

$$p_\theta(\theta_k) = \frac{1}{2\pi}(1 + \alpha \cos\theta_k). \tag{S2}$$

Calling X and Y the coordinates of the end point of the walk along the x and y axes, one has

$$X = \sum_{k=1}^{N} \cos\theta_k \quad \text{and} \quad Y = \sum_{k=1}^{N} \sin\theta_k. \tag{S3}$$

Using the PDF of θ_k from equation (S2), one gets

$$\langle X \rangle = \sum_{k=1}^{N} \int_0^{2\pi} \cos\theta_k \, p_\theta(\theta_k) \, d\theta_k = \frac{\alpha N}{2}, \tag{S4}$$

and similarly,

$$\langle Y \rangle = \sum_{k=1}^{N} \int_0^{2\pi} \sin\theta_k \, p_\theta(\theta_k) \, d\theta_k = 0. \tag{S5}$$

The second moments of X and Y can be computed by noting that the PDF of the random variables $\varphi_{kl} = \theta_k + \theta_l$ $(k \neq l)$ and $\psi_{kl} = \theta_k - \theta_l$ are identical and given by

$$p_\varphi(\varphi_{kl}) = p_\theta(\theta_k) * p_\theta(\theta_l) = \frac{1}{2\pi}\left(1 + \frac{\alpha^2}{2}\cos\varphi_{kl}\right). \tag{S6}$$

After derivation, this leads to

$$\langle X^2 \rangle = \int_0^{2\pi} \left(\sum_{k=1}^{N} \cos\theta_k\right)^2 p_\theta(\theta_k)\, d\theta_k = \frac{N}{2} + N(N-1)\frac{\alpha^2}{4}. \tag{S7}$$

The variance of X is then directly obtained

$$\mathrm{var}(X) = \sigma_X^2 = \langle X^2 \rangle - \langle X \rangle^2 = \frac{N}{2}\left(1 - \frac{\alpha^2}{2}\right). \tag{S8}$$

Similarly, one gets

$$\mathrm{var}(Y) = \langle Y^2 \rangle = N/2. \tag{S9}$$

Since we are in the configuration where $N \gg 1$, one can assume that both X and Y have normal distributions, which means are given by equations (S4) and (S5), and variances by equations (S8) and (S9).

From there, the expected value of the logarithm of the Schuster p-value is directly computed

$$\langle -\ln p \rangle = \frac{\langle X^2 \rangle + \langle Y^2 \rangle}{N} = 1 + \frac{N\alpha^2}{4}. \tag{S10}$$

In the case where there actually is some periodicity in the catalog at the studied period ($\alpha \neq 0$), the condition $N \gg 1$ might lead to $N\alpha^2 \gg 1$, and then the expected value of the Schuster p-value simply reduces to

$$\langle \ln p \rangle = -N\alpha^2/4. \tag{S11}$$

Computing the variance of $\ln p$ requires the tedious calculation of moments of order four: $\langle X^4 \rangle$, $\langle Y^4 \rangle$ and $\langle X^2 Y^2 \rangle$. Instead, one can notice that for small values of α (typically $\alpha < 0.5$), $\mathrm{var}(X) \approx \mathrm{var}(Y) = N/2$. For $\alpha \geq 0.5$ both variances become slightly different, but in that case, the Schuster walk mostly progresses in the x direction, and the variance of the end position of the walk is dominated by the variance of X itself. In order to compute the PDF of the distance D covered by the Schuster walk, one can thus assume that $\mathrm{var}(Y) = \sigma_X^2$

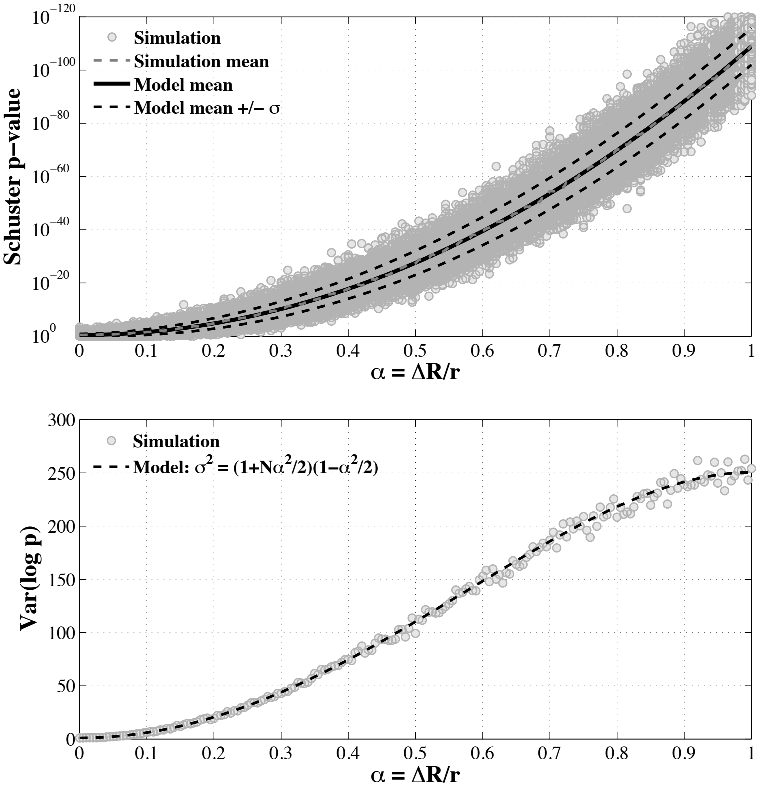


Figure S1: Comparison of simulated Schuster tests and theory for harmonic variations of seismicity rate of peak-to-peak amplitude $2\alpha = 2\Delta R/r$. Upper plot: Simulation for 100 values of α linearly distributed between 0 and 1. For each value of α, 200 catalogs of 1000 events each are simulated and the Schuster p-value is computed. The mean simulated p-value is plotted as a dashed gray line, while the mean predicted by the theory (equation (S10)) is plotted as a solid black line, and the dotted black line shows the 1-σ standard deviation (equation (S12)) above and below the mean. Lower plot: Comparison between the simulated and theoretical (equation (S12)) values of the standard deviation. The simulation is similar to the one in the upper plot, except that 2000 catalogs were drawn for each value of α.

for all values of α. Under this approximation, $D^2 = X^2 + Y^2$ has a noncentral chi-squared distribution with 2 degrees of freedom and a noncentrality parameter $\lambda = \langle X \rangle^2 / \sigma_X^2$. The properties of the noncentral chi-squared distribution are well known (*Muirhead*, 2005), and give us the variance of $\ln p$ in the case where $N \gg 1$:

$$\operatorname{var}(\ln p) = \left(1 - \frac{\alpha^2}{2}\right)\left(1 + \frac{N\alpha^2}{2}\right). \tag{S12}$$

Figure S1 shows that equations (S10) and (S12) represent good approximations of the mean and variance of the $\ln p$-value.

S.2 Computation of the complete PDF for ε_c

In order to derive the complete PDF for ε_c, we generated $N_s = 10^5$ uniformly distributed catalogs for each of which we computed a spectrum between $T_{\min}/t = 10^{-3}$ and $T_{\max}/t = 0.1$, extracted the minimum p-value δ_m and computed the value of the random variable $Y = \varepsilon N \delta_m$ with equation (2.12). We then compute the probability that the set of 10^5 random variables $\{Y_1, Y_2, \ldots, Y_{N_s}\}$ is an occurrence of the PDF in equation (2.14), for different values of ε_c.

S.2.1 Method

If a set of values occurs out of a uniform PDF over [0,1], we are able to analytically derive the PDF $p_s(s)$ of the standard deviation s of the binned values around the mean. This derivation is given in next subsection for clarity. With that in hand, here is the approach we follow: we first pick a value for ε_c, and we transform the set of values $Y = \{Y_1, Y_2, \ldots, Y_{N_s}\}$ into a set set of values $Z = \{Z_1, Z_2, \ldots, Z_{N_s}\}$, that would have a uniform distribution over [0,1] if equation (2.14) was the PDF for Y:

$$Z_i = \int_0^{Y_i} p_Y(y)\, dy = \int_0^{Y_i} e^{-y/\varepsilon_c} \frac{dy}{\varepsilon_c} = 1 - e^{-Y_i/\varepsilon_c}, \tag{S13}$$

We then bin the values of Z (i.e. we divide the interval [0,1] into b bins of equal size, and count how many values of Z fall in each bin), compute the standard deviation of the number of Z falling in each bin, and use this to compute the probability that the set of values Z occurs out of a uniform PDF. We repeat the process for different values of ε_c in order to derive the full PDF for ε_c.

S.2.2 Derivation of the PDF $p_s(s)$

In order to compute the probability that Z can be described by a uniform PDF, we divide the interval [0,1] into b bins of equal size and count the number n_k of Z_i that fall in bin number k ($k \in \{1, 2, \ldots, b\}$). The random variable n_k has thus the same binomial distribution

$\mathcal{B}(N_s, 1/b)$ in each bin:

$$p_n(n_k) = \binom{N_s}{n_k} \left(\frac{1}{b}\right)^{n_k} \left(1 - \frac{1}{b}\right)^{N_s - n_k}. \tag{S14}$$

Now, if N_s is large enough, this PDF can be approximated by a Gaussian distribution $n_k \sim \mathcal{N}\left(\frac{N_s}{b}, \frac{N_s}{b}\left(1 - \frac{1}{b}\right)\right)$. Finally, introducing the normalized and centered number of Z_i in bin number k

$$X_k = \frac{n_k}{N_s/b} - 1, \tag{S15}$$

X_k thus follows a centered normal distribution, independent of k

$$p_x(X_k) = \mathcal{N}\left(0, \frac{b-1}{N_s}\right) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}, \tag{S16}$$

where $\sigma^2 = \frac{b-1}{N_s}$.

In order to estimate the probability that a set of uniformly distributed variables has the standard deviation of $X = \{X_1, X_2, \ldots, X_b\}$, we use the unbiased estimator of the standard deviation of a set of b random variables with a centered Gaussian distribution

$$s = \frac{1}{\sqrt{b-1}} \sqrt{\sum_{i=1}^{b} X_k^2}. \tag{S17}$$

We can write

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} p_x(x_1)\, dx_1 \int_{-\infty}^{\infty} p_x(x_2)\, dx_2 \ldots \int_{-\infty}^{\infty} p_x(x_b)\, dx_b, \\ &= \int_{\mathbb{R}^b} p_x(x_1) p_x(x_2) \ldots p_x(x_b)\, dx_1\, dx_2 \ldots\, dx_b, \\ &= \int_{\mathbb{R}^b} \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^b e^{-\frac{x_1^2 + x_2^2 + \cdots + x_b^2}{2\sigma^2}}\, dx_1\, dx_2 \ldots\, dx_b. \end{aligned}$$

We do the spherical change of variable $r = \sqrt{x_1^2 + x_2^2 + \cdots + x_b^2}$, so we have $dx_1\, dx_2 \ldots\, dx_b =$

$S_b(r)\,dr$, where $S_b(r)$ is the surface of a sphere in b dimensions and is

$$S_b(r) = b\,\frac{\pi^{b/2}}{\Gamma\left(\frac{b}{2}+1\right)}\;r^{b-1},$$

where Γ is the usual Gamma function defined by

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}\,dt.$$

So we now have

$$\begin{aligned} 1 &= \int_0^\infty \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^b e^{-\frac{r^2}{2\sigma^2}}\,S_b(r)\,dr, \\ &= \int_0^\infty \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^b e^{-\frac{r^2}{2\sigma^2}}\,b\,\frac{\pi^{b/2}}{\Gamma\left(\frac{b}{2}+1\right)}\;r^{b-1}\,dr, \\ &= \int_0^\infty \frac{b}{\Gamma\left(\frac{b}{2}+1\right)}\left(\frac{1}{\sigma\sqrt{2}}\right)^b r^{b-1}e^{-\frac{r^2}{2\sigma^2}}\,dr. \end{aligned}$$

We finally introduce the unbiased estimator of the standard deviation as our new variable

$$s = \frac{r}{\sqrt{b-1}}, \qquad dr = ds\sqrt{b-1},$$

which finally gives us

$$\begin{aligned} 1 &= \int_0^\infty \frac{b}{\Gamma\left(\frac{b}{2}+1\right)}\left(\frac{1}{\sigma\sqrt{2}}\right)^b \sqrt{b-1}^{\,b-1} s^{b-1} e^{-\frac{\sqrt{b-1}^2 s^2}{2\sigma^2}}\;\sqrt{b-1}ds, \\ &= \int_0^\infty \frac{b}{\Gamma\left(\frac{b}{2}+1\right)}\left(\frac{1}{\tilde{\sigma}\sqrt{2}}\right)^b s^{b-1}e^{-\frac{s^2}{2\tilde{\sigma}^2}}\,ds, = \int_0^\infty p_s(s)\,ds, \end{aligned}$$

where

$$\tilde{\sigma} = \frac{\sigma}{\sqrt{b-1}} = \frac{1}{\sqrt{b-1}}\sqrt{\frac{b-1}{N_s}} = \frac{1}{\sqrt{N_s}},$$

is actually independent of b.

Finally, using the property of the Gamma function $\Gamma(x+1) = x\Gamma(x)$, the PDF for the standard deviation of X_k is

$$p_s(s) = \frac{2}{\Gamma\left(\frac{b}{2}\right)} \sqrt{\frac{N_s}{2}}^{\,b} s^{b-1} e^{-\frac{N_s s^2}{2}} \tag{S18}$$

S.3 Generation of earthquake catalogs following a given seismicity rate

The spectra on figure 2.4 have been computed from earthquake catalogs randomly generated from a specific seismicity rate $R(t_k)$, where hereafter, t_k is the time of earthquake number k. Taking the total length of the catalog to be the time unit and

$$N = \int_0^1 R(\tau)\, d\tau \tag{S19}$$

the total number of events in the catalog, one can define the probability of an event to happen at time t_k:

$$p_R(t_k) = \frac{R(t_k)}{N}. \tag{S20}$$

In order to generate a catalog that follows the seismicity rate R, we thus generate N times $t^u = \{t_1^u, t_2^u, \ldots, t_N^u\}$ uniformly distributed over [0,1], and associate to each of them a time of event t_k solution of the equation

$$t_k^u = \int_0^{t_k} p_R(\tau)\, d\tau. \tag{S21}$$

The set of times t_k is thus a realization of the PDF p_R.

S.3.1 Harmonic seismicity rate: figure 2.4a

In order to generate a catalog that contains N events occurring out of a harmonically varying seismicity rate (equation (2.3) in the main paper):

$$\frac{R(t_k)}{r} = 1 + \alpha \cos \frac{2\pi t_k}{T}, \tag{S22}$$

the events time distribution will follow the PDF

$$p_R(t_k) = 1 + \alpha \cos \frac{2\pi t_k}{T}. \tag{S23}$$

From the uniformly distributed times t^u, the times of events t_k are solution of equation (S21), i.e.

$$t_k^u = t_k + \alpha \frac{T}{2\pi} \sin \frac{2\pi t_k}{T}, \tag{S24}$$

which can be efficiently solved by a Newton-Raphton (*Press et al.*, 1992) algorithm. Noting

$$\theta_u = \frac{2\pi}{T} \begin{pmatrix} t_1^u \\ t_2^u \\ \vdots \\ t_N^u \end{pmatrix}, \quad \text{and} \quad \theta_e = \frac{2\pi}{T} \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_N \end{pmatrix}, \tag{S25}$$

equation (S24) simply becomes

$$\theta_u = \theta_e + \alpha \sin \theta_e. \tag{S26}$$

Introducing the function to zero out $f(\theta) = \theta + \alpha \sin \theta - \theta_u$, starting with $\theta_0 = \theta_u$, one has

$$\theta_{n+1} = \theta_n - \frac{f(\theta_n)}{f'(\theta_n)} = \theta_n - \frac{\theta_n + \alpha \sin \theta_n - \theta_u}{1 + \alpha \cos \theta_n}, \tag{S27}$$

and we take $\theta_e = \theta_n$ such that $\|\theta_n - \theta_{n-1}\| < \epsilon$, where $\epsilon \ll 1$. In our algorithm, we took $\epsilon = 10^{-5}$. Since the θs here are vectors, the division sign is abusive, and actually refers to a term to term division of each vector. The first term of the series θ_1 is actually very simple, and we thus directly started the algorithm with it:

$$\theta_1 = \theta_u - \frac{\alpha \sin \theta_u}{1 + \alpha \cos \theta_u}.$$

S.3.2 Periodic non-harmonic seismicity rate: figure 2.4b

The spectrum in figure 2.4b has been computed from an earthquake catalog following the PDF

$$p_R(t_k) = \frac{1}{\Omega} \begin{cases} 1, & \text{if} \quad t_k[T]/T \in [0; 0.1] \\ \alpha, & \text{if} \quad t_k[T]/T \in [0.1; 0.2] \\ 1, & \text{if} \quad t_k[T]/T \in [0.2; 1] \end{cases}, \tag{S28}$$

where the normalization factor $\Omega = 0.9{+}0.1\alpha$, and $t_k[T]$ is the modulus of t_k after division by the period T. In this case, the computation of t_k from t_k^u is quite tedious but straightforward. We here directly present the result:

$$t_k = t_k^u + \begin{cases} 0.1 t_k^u[T](\alpha - 1), & \text{if } t_k^u[T]/T \in [0; \frac{0.1}{\Omega}] \\ (1 - \frac{1}{\alpha})(0.1T - 0.9 t_k^u[T]), & \text{if } t_k^u[T]/T \in [\frac{0.1}{\Omega}; \frac{0.1(1+\alpha)}{\Omega}] \\ 0.1(\alpha - 1)(t_k^u[T] - T), & \text{if } t_k^u[T]/T \in [\frac{0.1(1+\alpha)}{\Omega}; 1] \end{cases}. \tag{S29}$$

The catalog for which the spectrum is plotted in figure 2.4b was generated with $\alpha = 3$.

S.3.3 Aftershock sequence: figure 2.4c

In order to generate an aftershock sequence, we choose a seismicity rate following the law proposed by *Dieterich* (1994):

$$R(t_k) = \frac{r}{1 + \left(e^{-\Delta\tau/A\sigma} - 1\right) e^{-(t_k - t_m)/t_a} \mathcal{H}(t_k - t_m)}, \tag{S30}$$

where r is the background seismicity rate supposed to be identical before and after the aftershock sequence, $\Delta\tau$ is the amplitude of the stress step caused by the mainshock generating the aftershock sequence, A is a dimensionless fault parameter, σ is the normal stress on the fault, t_m is the time of the mainshock and t_a the characteristic duration of the aftershock sequence. In order to make sure that the aftershock sequence is over before the end of the catalog, we suppose that $t_a \ll 1 - t_m$ (recall that the total duration of the catalog is $t = 1$). The function $\mathcal{H}(.)$ is the Heavyside function ($\mathcal{H}(x) = 0$ for $x < 0$ and $\mathcal{H}(x) = 1$ for $x \geq 0$). Note that individual values of parameters $\Delta\tau$, A and σ actually don't matter, the result depending only on the value of the ratio

$$Q = \Delta\tau / A\sigma. \tag{S31}$$

The seismicity rate from equation (S30) can be separated into a constant background rate r and the seismicity rate of the aftershock sequence $R_a(t)$:

$$R(t_k) = r + \frac{r\mathcal{H}(t_k - t_m)}{\frac{1}{1-e^{-Q}} e^{-(t_k - t_m)/t_a} - 1} = r + R_a(t_k). \tag{S32}$$

The number of events in the aftershock sequence being

$$N_a = \int_0^1 R_a(\tau)\, d\tau \approx r t_a \frac{\Delta\tau}{A\sigma} = r t_a Q, \tag{S33}$$

we generate the complete catalog by both generating r events uniformly distributed over [0,1] and N_a events occurring out of the PDF $p_{Ra}(t_k) = R_a(t_k)/N_a$. More specifically, from N_a times t^u uniformly distributed over [0,1], the times of events $t_e = \{t_1, t_2, \ldots, t_N\}$ in the aftershock sequence can be derived from equation (S21):

$$t_e = t_m - t_a \ln \frac{e^Q - e^{Qt^u}}{e^Q - 1}. \tag{S34}$$

S.4 Seismicity in Nepal

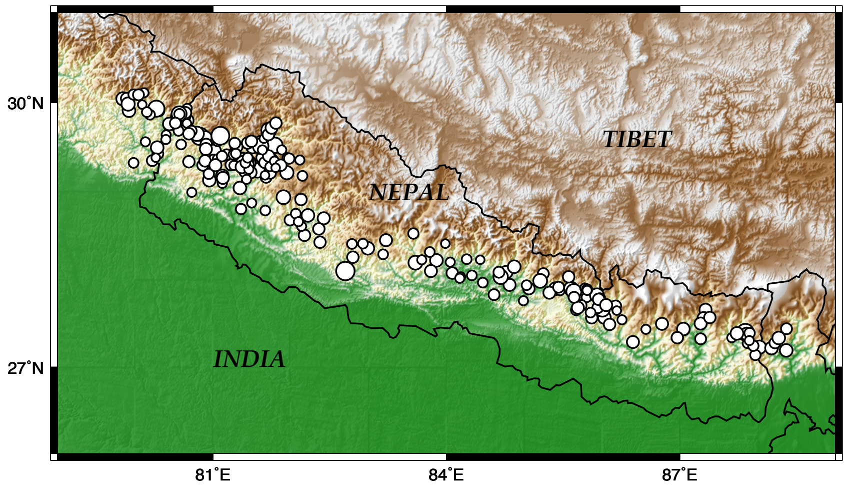


Figure S2: Maps showing $M_b \geq 4$ midcrustal events from the ISC catalog used in the study, selected according to their position on the map, using the same selection contour as in *Bollinger et al.* (2007). Circles sizes are proportional to events magnitudes: same scale for event sizes as on the NSC map.

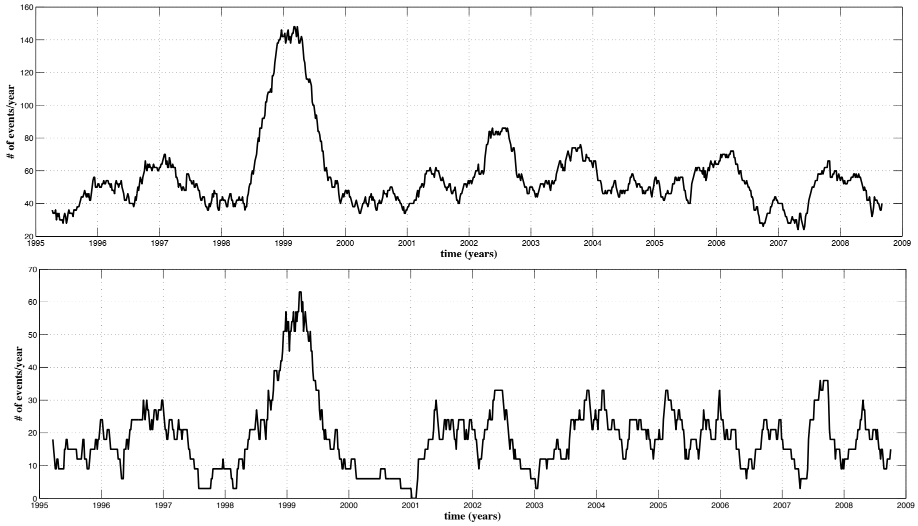


Figure S3: Midcrustal seismicity rate from the NSC catalog in Nepal, from 1995 to 2008, for events of local magnitude $M_L \geq 3.5$ (upper plot), and $M_L \geq 4$ (lower plot). The seismicity rate has been computed using a sliding window of half a year, centered on the date at which the seismicity rate is evaluated. A clear increase of seismicity rate appears during winter 1998-1999.

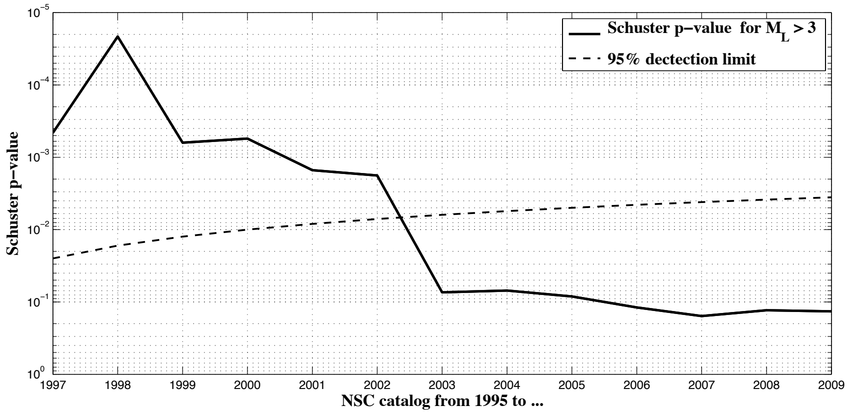


Figure S4: Evolution of the Schuster p-value at 1 year by gradually adding years to the NSC $M_L \geq 3$ declustered catalogs, starting in 1995. The dashed black line indicates detection at the 95% confidence level.

Chapter 3

The role of velocity-neutral creep on the modulation of tectonic tremor activity by periodic loading

Published in Geophysical Research Letters, VOL. 39, L16310, doi:10.1029/2012GL052326, 2012

Thomas Ader,[1,2] Jean-Paul Ampuero,[1] Jean-Philippe Avouac[1]

[1]Department of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA 91125, USA.

[2]Laboratoire de Géologie, Ecole Normale Supérieure, CNRS, 24 rue Lhomond, 75004 Paris, France.

Abstract

Slow slip events and associated non-volcanic tremors are sensitive to oscillatory stress perturbations, such as those induced by tides or seismic surface waves. Slow slip events and tremors are thought to occur near the seismic-aseismic transition regions of active faults, where the difference $a - b = \partial\mu/\partial \ln V$ between the sensitivity of friction to slip rate and fault state, which characterizes the stability of slip, can be arbitrarily small. We investigate the response of a velocity-strengthening fault region to oscillatory loads through analytical approximations and spring-slider simulations. We find that fault areas that are near velocity-neutral at steady-state, i.e., $\partial\mu/\partial \ln V \approx 0$, are highly sensitive to cyclic loads: oscillatory stress perturbations in a certain range of periods induce large transient slip velocities. These aseismic transients can in turn trigger tremor activity with enhanced oscillatory modulation. In this sensitive regime, a harmonic Coulomb stress perturbation of amplitude ΔS causes a slip rate perturbation varying as $e^{\Delta S/(a-b)\sigma}$, where σ is the effective normal stress. This result is in agreement with observations of the relationship between tremor rate and amplitude of stress perturbations for tremors triggered by passing seismic waves. Our model of tremor modulation mediated by transient creep does not require extremely high pore fluid pressure and provides a framework to interpret the sensitivity and phase of tidally modulated tremors observed in Parkfield and Shikoku in terms of spatial variations of friction parameters and background slip rate.

3.1 Introduction

The recent discovery of slow-slip events (SSEs) and non-volcanic tremors (NVTs) has lead to a vast body of observational work in the past decade. SSEs and NVTs appear to coincide in time and space (e.g., *Rogers and Dragert* (2003)) and have been observed in various tectonic settings (*Schwartz and Rokosky* (2007); *Brown et al.* (2009); *Shelly et al.* (2011)). *Rubinstein et al.* (2008), *Nakata et al.* (2008) and *Thomas et al.* (2009, 2012) reported a modulation of NVTs intensity by tidal stresses of a few kPa or less, in the Cascadia subduction zone, in southwest Japan and on the deep San Andreas Fault at Parkfield, respectively. *Hawthorne and Rubin* (2010) inferred a modulation of the slip rate of SSEs by tidal stresses studying borehole strainmeter data in Cascadia. *Miyazawa and Brodsky* (2008) found that NVTs in western Japan were triggered by the passing surface waves radiated by the 2004 Sumatra earthquake, and observed an exponential relationship between the amplitude of the NVTs and the Coulomb stress perturbation in the source region.

Nakata et al. (2008) and *Thomas et al.* (2009, 2012) explained the correlation of NVTs with tidal loading with the model of *Dieterich* (1994), therefore postulating that tidal stresses directly triggered seismic slip on locked asperities. This interpretation required low values of $a\sigma$, either implying a orders of magnitude lower than values inferred from lab experiments (*Blanpied et al.*, 1995) or extremely low effective normal stresses. They retained the second hypothesis (*Nakata et al.* (2008) proposed $\sigma_{\text{eff}} \approx 100$ kPa while *Thomas et al.* (2009) found $\sigma_{\text{eff}} \approx 9$ to 35 kPa), and justified it by a nearly lithostatic pore pressure. However, since the nucleation size on an asperity is inversely proportional to the effective normal stress (e.g., *Rubin and Ampuero* (2005)), this would imply large nucleation sizes for seismic ruptures, in contradiction with the prevailing view that tremors are small shear rupture events. *Miyazawa and Brodsky* (2008) explained the exponential relationship between tremor amplitude variations and amplitude of the incoming waves with a pre-existing exponential distribution of failure stresses within the tremor source region.

Velocity-weakening fault patches also show enhanced sensitivity to oscillatory loads, but only over a narrow range of patch sizes and loading periods (*Perfettini and Schmittbuhl*, 2001; *Lowry*, 2006).

Here we present an alternative mechanism for these observed correlations, relying on the fact that tremors usually happen at the transition between the rate-strengthening and rate-weakening parts of a fault, thus a region where the sensitivity of steady-state friction to velocity, $a - b = \partial\mu/\partial \ln V$, can be arbitrarily low. We first present the response of a spring-slider system with rate-strengthening rheology to harmonic shear and normal stress perturbations of different periods. We then establish an non-linear, exponential relationship between the amplitudes of the slip rate and the stress perturbations for large enough Coulomb stress perturbations.

3.2 Model hypotheses

We adopt the view proposed by *Ide et al.* (2007) and *Shelly et al.* (2011) that tremors are generated by the rupture of small rate-weakening asperities caused by slip on the surrounding plate interface. Under the assumption that the rupture is Coulombian, the NVTs intensity is directly proportional to the slip rate on the fault.

We thus study the response of a rate-strengthening fault to a stress perturbation, modeling the fault as a one-dimensional spring-slider system with stiffness k (e.g., *Perfettini et al.*, 2001; *Parsons*, 2004), loaded at constant background velocity V_{ss}, under shear stress perturbation $\Delta\tau(t) = \Delta\tau e^{i\omega t}$ and perturbed normal stress $\sigma(t) = \sigma_o + \Delta\sigma e^{i\omega t}$, where the amplitude of the perturbation is smaller than the prevailing normal stress ($\Delta\sigma < \sigma_o$). Both shear and normal stress perturbations are supposed to be in phase for the sake of simplicity. The evolution of the friction coefficient μ is described by a rate-and-state law (e.g., *Marone*, 1998):

$$\mu = \mu_{\text{ss}} + a \ln \frac{V}{V_{\text{ss}}} + b \ln \frac{\theta V_{\text{ss}}}{D_c}, \tag{3.1}$$

where V is the total slip rate of the slider, θ a fault state variable, μ_{ss} the steady-state friction coefficient at slip rate V_{ss}, D_c the characteristic slip for friction to evolve between two steady states, and a and b are constitutive fault parameters verifying $a - b > 0$, such that the system has a rate-strengthening rheology. The state variable θ evolves according to the "aging law" (e.g., *Marone*, 1998):

$$\frac{d\theta}{dt} = 1 - \frac{V\theta}{D_c}. \tag{3.2}$$

3.3 Period dependent response of the system

When the amplitude of the harmonic perturbations of Coulomb stress $\Delta S = \Delta\tau - \mu_{\text{ss}}\Delta\sigma$ is small enough (i.e., $\Delta S \ll (a-b)\sigma$), the slip rate of the slider undergoes small harmonic variations around its steady state value: $V(t) = V_{\text{ss}} + \Delta V e^{i\omega t}$, where $\Delta V \ll V_{\text{ss}}$. The resulting perturbations of slip rate ΔV can be obtained by a linearized approximation

(*Segall*, 2010):

$$\frac{\Delta V}{V_{\text{ss}}} = \frac{i\omega}{1 + i\omega t_a(\omega)} \frac{\Delta S}{\dot{\tau}_{\text{ss}}}, \tag{3.3}$$

where $\omega = 2\pi/T$ is the pulsation, $\dot{\tau}_{\text{ss}} = kV_{\text{ss}}$ the background stressing rate, $t_a(\omega) = A(\omega)\sigma/\tau_{\text{ss}}$, and

$$A(\omega) = a - \frac{b}{1 + i\omega\theta_{\text{ss}}} \tag{3.4}$$

is a period dependent constitutive fault parameter. In the limits $\omega\theta_{\text{ss}} \ll 1$ and $\omega\theta_{\text{ss}} \gg 1$, A becomes real ($A = a - b$ and $A = a$, respectively) and quantifies the velocity dependence of the friction in the steady-state regime.

Equation (3.3) is represented for two different values of ΔS in Figure 3.1 (dashed line with triangles), such that $\Delta S_1 < (a-b)\sigma$ and $\Delta S_2 > (a-b)\sigma$. Three characteristic periods bounding different behaviors of the system appear. $T_\theta = 2\pi\theta_{\text{ss}} = 2\pi D_c/V_{\text{ss}}$, where θ_{ss} is the steady-state value of the state variable, defines the characteristic time scale for the evolution of the state variable. For perturbations with period $T < T_\theta$, the state variable does not have time to evolve and the rate-and-state law reduces to a purely rate-dependent law with $\partial\mu/\partial \ln V = a$. $T_Q = T_\theta \times a/(a-b) > T_\theta$ is the period above which the state variable has time to fully adjust so that in the steady state, $\partial\mu/\partial \ln V = a - b$. The third period, T_a, is the one for which $|\omega t_a(\omega)| = 1$ and separates between two physically different responses of the system to the stress perturbation. For periods $T > T_a$, the damping due to the friction acts on a much smaller time-scale than that of the characteristic evolution of the spring-slider and the response of the system becomes that of a perturbed spring-slider in steady state with no friction. In this quasi-static regime, equation (3.3) reduces to:

$$\frac{\Delta V}{V_{\text{ss}}} = \frac{\Delta \dot{S}}{\dot{\tau}_{\text{ss}}}. \tag{3.5}$$

For $T < T_a$, the period of the velocity oscillations is too small for the spring stiffness to have any significant effect, and the system evolves as a simple slider with a rate-and-state friction law. In this regime, the amplitude of the velocity perturbation is proportional to the amplitude of the stress perturbation:

$$\frac{\Delta V}{V_{\text{ss}}} = \frac{\Delta S}{A\sigma}, \tag{3.6}$$

where $A = |A(\omega)|$ depends on the period. In particular, when $T_Q < T$, $A \approx (a-b)$ and so

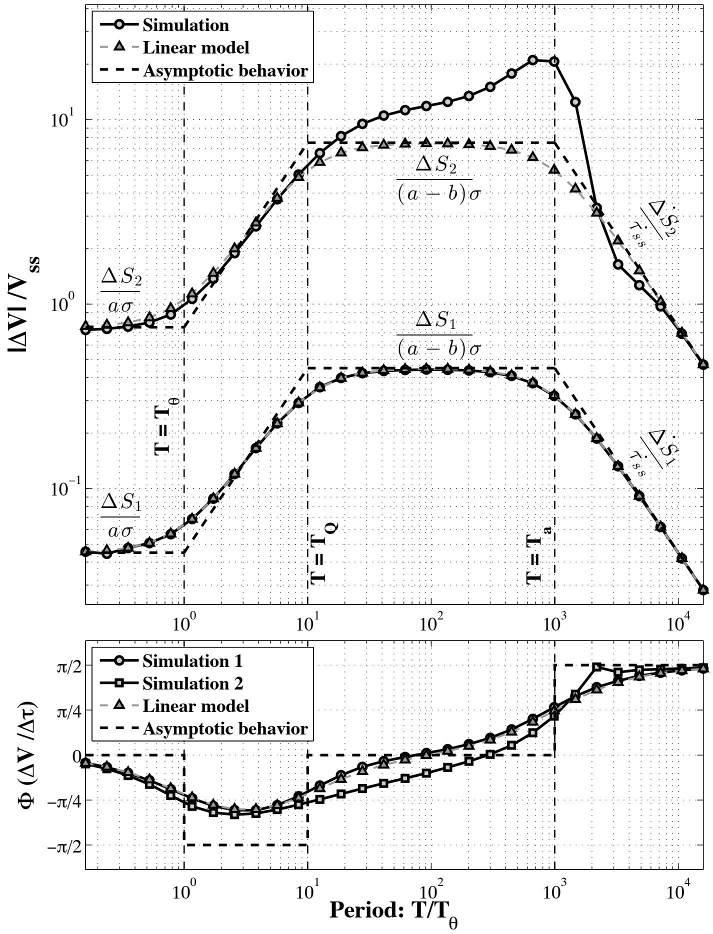


Figure 3.1: Response of a spring-slider system to small harmonic Coulomb stress perturbations of different periods and amplitudes $\Delta S_1 = 0.9$ kPa (simulation 1: $\Delta\tau_1 = \Delta\sigma_1 = 3$ kPa) and $\Delta S_2 = 15$ kPa (simulation 2: $\Delta\tau_2 = \Delta\sigma_2 = 50$ kPa). The system is undergoing constant loading at velocity $V_{\text{ss}} = 0.02$ m/yr under mean normal stress $\sigma_o = 5$ MPa. The normalized spring stiffness is $k/\sigma_o = 0.002$ m^{-1}. The other parameters are: $\mu_{\text{ss}} = 0.7$, $a = 0.004$, $b = 0.0036$ and $D_c = 0.2$ mm. Upper panel: Amplitude of the creep rate variations. The black lines with circles represents the results of the simulations (one circle for each period tested). The dashed grey lines with triangles represent the small perturbation approximation (equation (3.3)) for each simulation while the dashed black lines indicate the corresponding asymptotic behavior of the system with equations indicated on the plot. The critical periods T_θ, T_Q and T_a are also indicated on the plot. Lower panel: Phase difference between the creep rate and the Coulomb stress variations.

the amplitude of the velocity oscillations becomes:

$$\frac{\Delta V}{V_{\text{ss}}} = \frac{\Delta S}{(a-b)\sigma}, \tag{3.7}$$

which may result in large oscillations of the slip rate for small values of $(a - b)$.

The ratio T_a/T_θ has the following expression:

$$\left(\frac{T_a}{T_\theta}\right)^2 = \frac{1}{2}\left[\sqrt{4\tilde{a}^2 + \left(1 - \left[\tilde{a} - \tilde{b}\right]^2\right)^2} - \left(1 - \left[\tilde{a} - \tilde{b}\right]^2\right)\right], \tag{3.8}$$

where $\tilde{a} = a\sigma/kD_c$ and $\tilde{b} = b\sigma/kD_c$. A graphical representation of equation (3.8) is given in the supplementary figure S1. A response following equation (3.7) requires $T_a/T_\theta \gg 1$, which is possible only when $\tilde{a} - \tilde{b} \gg 1$, in which case $T_a/T_\theta = \tilde{a} - \tilde{b}$.

In order to assess the validity of the linear approximation, we simulate the general response of a spring-slider system to harmonic shear and normal stress perturbations of equal amplitudes, solving the equations of motion using a Runge-Kutta algorithm with a fifth-order adaptive step-size control (*Press et al.*, 1992) for Coulomb stress perturbations of amplitudes ΔS_1 and ΔS_2. The results are plotted on Figure 3.1 and show that in the first case, the linear approximation is justified, while in the second case, for periods $T_Q < T < T_a$, the amplitude of the slip perturbation becomes non linear and greater than what equation (3.3) predicts.

3.4 Influence of Coulomb stress amplitude

Over the range of periods $T_Q < T < T_a$, Coulomb stress perturbations of amplitude greater than $(a - b)\sigma$ induce non linear velocity fluctuations of large amplitude. In Figure 3.1, $T_a/T_Q = (a-b)/a \times T_a/T_\theta = 10^2$, but can actually be several orders of magnitude larger for a different set of parameters values. For instance, $D_c = 2\mu$m (e.g., *Marone*, 1998) increases this ratio to 10^4. This non-linear amplification of the response can thus prevail over a range of periods spanning several orders of magnitude.

To first approximation, the induced non linear velocity fluctuations depend exponentially on the stress perturbation (see derivation in appendix):

$$\frac{\Delta V}{V_{\rm ss}} \approx e^{\frac{\Delta S}{(a-b)\sigma}}. \tag{3.9}$$

Equation (3.7) is simply a linear approximation of equation (3.9) when $\Delta S \ll (a - b)\sigma$. Figure 3.2 shows the result of a simulation with shear and normal stress perturbations of period $T/T_Q = T_a/T > 1$, and increasing amplitudes. This simulation shows that the exponential approximation (equation (3.9)) provides a good description of the system's behavior. This statement still holds when considering the “slip law” (e.g., *Marone*, 1998) instead of the “aging law” for the evolution of the state variable (supplementary figure S2).

This simulation also predicts a correlation of the slip perturbation with the Coulomb stress rather than with the shear stress perturbation. This point has been discussed in the observational literature, but unfortunately the conclusions remain elusive and thus hard to compare with our model predictions. Both *Nakata et al.* (2008) and *Miyazawa and Brodsky* (2008) reported a correlation of the tremor with the Coulomb stress perturbations, while *Thomas et al.* (2009, 2012) and *Hawthorne and Rubin* (2010) observed a correlation with shear stress variations only. However, *Thomas et al.* (2009, 2012) found the best correlation for an extremely small friction coefficient ($\mu = 0.02$), while *Hawthorne and Rubin* (2010) noted that if fluids did not diffuse significantly over the time scale of tides, the changes in pore pressure could compensate the applied normal stress variations, resulting in small effective normal stress variations. In both cases, the effective Coulomb stress and shear stress variations were almost the same, making it impossible to ascertain whether tremors correlated better with the one or the other.

Looking at the phase difference $\Phi(\Delta V/\Delta\tau)$ between the slip rate and the stress variations (Figure 3.2, lower plot) indicates that, at periods for which the sensitivity is the highest, NVTs should correlate with stress perturbations rather than with perturbations of the stress rate ($\Phi = 0$ and not $\pi/2$). We will come back to this point in the discussion.

3.5 Discussion and Conclusions

We here propose a mechanism to explain the observed triggering of NVTs by tidal stresses and passing seismic surface waves (*Miyazawa and Brodsky*, 2008; *Nakata et al.*, 2008), as well as the apparent tidal modulation of slow slip in Cascadia (*Hawthorne and Rubin*, 2010). The idea relies on the fact that both NVTs and SSEs seem to occur right below the locked section of faults, where the fault constitutive parameters define a nearly velocity-neutral zone ($a - b \approx 0$). We show that for a certain range of periods, a harmonic perturbation of the Coulomb stress on such a fault can induce a large perturbation of the slip rate around its steady-state value, of amplitude varying exponentially with the amplitude of the Coulomb stress perturbation. Assuming that NVTs are due to the rupture of rate-weakening ($a-b < 0$) patches embedded in that fault region, the tremor intensity should be proportional to the transient aseismic slip velocity. This can explain the sensitivity of NVTs to tidal and seismic stresses without requiring unusual values for a and b nor requiring extremely low

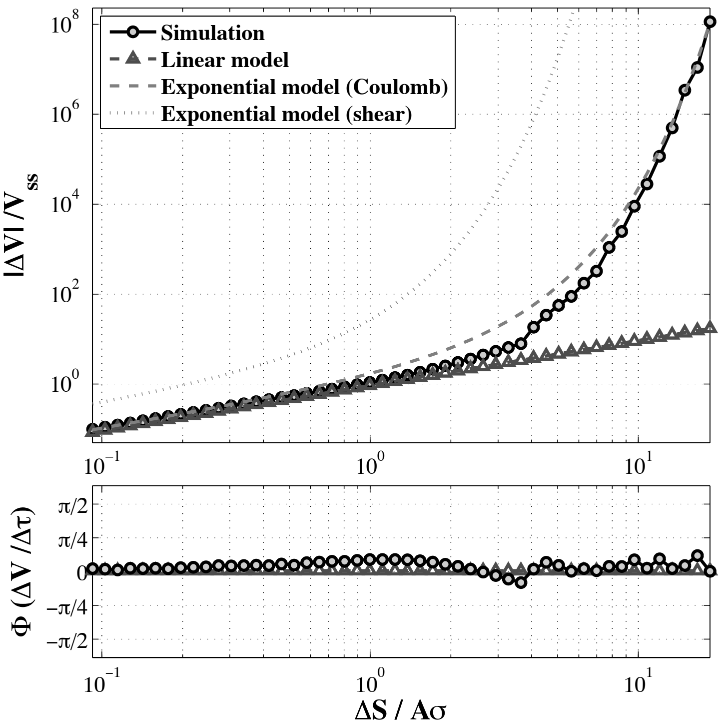


Figure 3.2: Non linear response of a spring-slider system to small harmonic stress perturbations for different amplitudes. The period T of the perturbation is such that $T/T_Q = T_a/T = 2.5$. The parameters are the same as in Figure 3.1, except for the fault parameter $b = 0.00385$ and $D_c = 0.5$ mm, so that $A = |A(\omega)| = 1.08(a - b) \approx (a - b)$. The meaning of the different lines is given in the legend. Upper panel: Amplitude of the creep rate variations. The exponential models derive from equation (3.9) taking either the Coulomb stress or only the shear stress, and replacing $(a - b)$ by the actual value of A. The linear model corresponds to equation (3.3). The lower panel shows the phase difference between the creep rate and the Coulomb stress variations.

effective normal stresses. This also provides an alternative explanation to the exponential relationship between NVTs intensity and passing surface waves amplitude without resorting to ad-hoc exponential distributions of initial stresses (*Miyazawa and Brodsky*, 2008).

This model predicts a correlation of NVTs with varying stresses only for a bounded range of periods, $T_Q < T < T_a$, which can vary in space and time yielding inhomogeneities of the sensitivity of NVTs to stress perturbations. In Parkfield, *Thomas et al.* (2012) observed an increase of sensitivity at tidal periods as a function of depth and closeness to the creeping segment, i.e., towards regions where V_{ss} is expected to be larger. Given that both T_Q and T_a are inversely proportional to V_{ss}, those observations can be explained by the fact that as V_{ss} decreases, T_Q increases to values that might become higher than the tidal period, thus inhibiting the correlation. Other parameters, such as D_c, a and b, might also induce spatial variations of the sensitivity by acting on the bounding periods. Inhomogeneities of $(a-b)$ have two effects: they affect the bounding periods and they directly influence the amplitude of the correlation, according to equation (3.9). The existence of NVTs itself, and thus of rate-weakening patches $(a-b<0)$ within the creeping zone $(a-b \gtrsim 0)$, may stand as a direct manifestation of the non-uniformity of $(a-b)$ on the fault. This could explain variations of sensitivity to tides in Shikoku (*Ide*, 2010), which pattern seem too erratic to be explained by local variations of the creep rate. Finally, the analysis of the phase suggests that NVTs should correlate and approximately be in phase with the perturbing stress (Figure 3.2). However, Figure 3.1 shows that should the period of the perturbation get closer to T_Q or T_a, the maximum NVTs intensity would respectively happen slightly before or after the maximum stress perturbation. This could explain why *Nakata et al.* (2008) observed a time advance of tremors relative to tidal stresses (and concluded of a correlation with the stress rate coupled with a delayed nucleation), while *Thomas et al.* (2012) reported a slight time lag. In the latter study, if our interpretation of the loss of sensitivity due to decreasing V_{ss} is correct, the loss of correlation should go hand in hand with an increase of the phase lag.

Stress perturbations due to either tides or passing surface waves have been reported to be of the order of a few kPa (*Miyazawa and Brodsky*, 2008; *Nakata et al.*, 2008; *Thomas et al.*, 2012). For a hydrostatic effective normal stress of 300 MPa and stress perturbations of ampltiude 3 kPa, this exponential regime would be observed for $0 < a - b < 10^{-5}$. Although such values might appear small, they may prevail at the transition between the rate-weakening and rate-strengthening parts of the fault, where NVTs and SSEs are observed to originate. For $(a-b)/a = 10^{-2}$, $V_{\mathrm{ss}} \sim 1$ m/yr (e.g., *Schwartz and Rokosky*, 2007) and $D_c \sim 3\ \mu$m (e.g., *Marone*, 1998), T_Q is of the order of half a day, the dominant period of tides. One needs smaller D_c or larger V_{ss} for T_Q to become of the order of the seismic waves period.

In order to apply this mechanism to real faults, small values of $(a-b) = \partial\mu/\partial \ln V$ are necessary over large enough regions, in order to sustain the high sensitivity of tremors to oscillatory stresses. *Shimamoto* (1986) and *Moore et al.* (1997) reported a N-shaped dependence of the steady-state friction on $\ln V_{\mathrm{ss}}$ during lab experiments on halite and chyrsotile serpentine respectively, thus unraveling two critical velocities for which $\partial\mu/\partial \ln V = 0$. *Estrin and Bréchet* (1996) and more recently *Beeler* (2009) proposed models for frictional sliding with a N-shaped curve for the velocity dependence of the friction coefficient. *Shibazaki and Iio* (2003) and *Shibazaki and Shimamoto* (2007) subsequently used similar friction laws in simulations of slow-slip events. In those models, the creep velocity during the SSE was such that $\partial\mu/\partial \ln V \approx 0$. The previous spring-slider analysis can then be applied to this configuration, and the spatial extent of the zone with small values of $a-b$ is as large as the region of active slow slip. This altogether qualitatively reconciles the various observations of correlation of slow slip events and associated non volcanic tremors to stress perturbations induced by tides and passing seismic waves, with fault parameters in agreement with laboratory values, no drastic constraints on local pore pressures, or initial distribution of stresses.

Although the present study focuses on tremors and SSEs, this mechanism might also be applied to regular earthquakes in some situations. In Nepal for instance, *Bollinger et al.* (2004) reported annual modulations of the seismicity, which turned out to be linked to small stress perturbations of a few kPa, due to varying surface loads caused by the local hydrological cycle (*Bettinelli et al.*, 2008). Modulation of seismicity by daily tides of similar stress amplitudes was not found. Given that the correlating seismicity forms a belt falling at the transition between the locked and creeping zones of the fault (*Ader et al.*, 2012a), the seasonal variations of the seismicity rate might be related to the mechanism proposed here. The lack of sensitivity to daily loadings is explained by our model if the period T_Q lies between one day and one year.

3.A Appendix: Coulomb stress perturbation of large amplitude

If $(a-b)$ is small, the small perturbation hypothesis might not be valid anymore. In the steady-state regime, for periods such that $T_Q = aT_\theta/(a-b) < T < T_a$, the equation of motion of the system becomes:

$$\Delta\tau e^{i\omega t} = \dot{\tau}_{\text{ss}}\left[\frac{\delta(t)}{V_{\text{ss}}} - t\right] + \left(\sigma_o + \Delta\sigma e^{i\omega t}\right)\left[\mu_{\text{ss}} + A\ln\frac{V(t)}{V_{\text{ss}}}\right], \tag{3.A.1}$$

where $A = |A(\omega)| \approx (a-b)$. The two terms in the right-hand side of the equation represent respectively the elastic stress due to the spring and the friction on the slider. Now, if the slider reaches high velocities but over a short time, its overall displacement remains small and the elastic force will have little impact on the system. The friction then dictates the evolution of the slider, and equation (3.A.1) reduces to:

$$(\Delta\tau - \mu_{\text{ss}}\Delta\sigma) = A\sigma \ln \frac{\Delta V}{V_{\text{ss}}}, \tag{3.A.2}$$

which leads to the following relation between the velocity and Coulomb stress perturbations:

$$\frac{\Delta V}{V_{\text{ss}}} = e^{\frac{\Delta S}{A\sigma}}. \tag{3.A.3}$$

Acknowledgments

This research was supported by NSF grant EAR-1015698, the Gordon and Betty Moore Foundation and the Southern California Earthquake Center (funded by NSF Cooperative Agreement EAR-0106924, NSF grant EAR0838495 and USGS Cooperative Agreement 02HQAG0008). We thank Paul Segall and another anonymous reviewer for useful comments. This paper is Caltech Tectonics Observatory contribution 206, Caltech Seismolab contribution 10081 and SCEC contribution 1647.

Supplementary figures

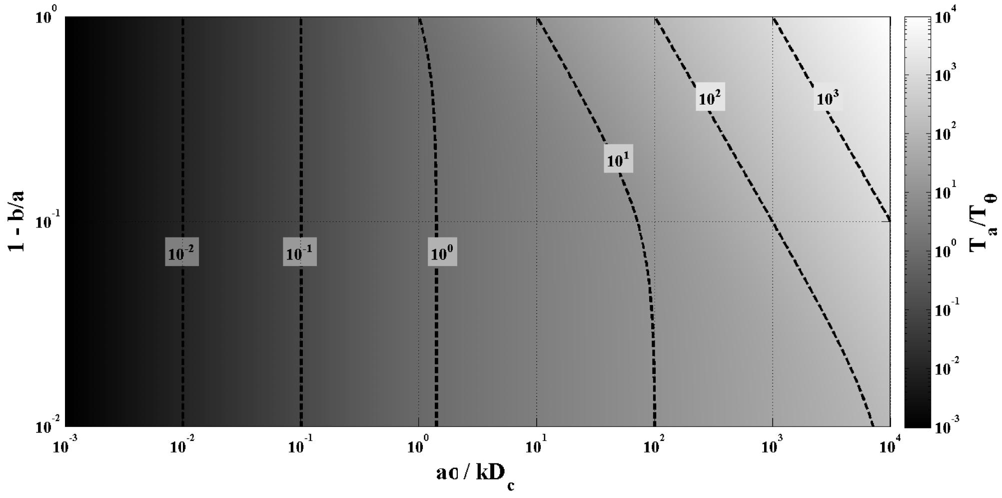


Figure S1: A graphical representation of equation (3.8) and gives the value of the ratio T_a/T_θ for different values of the problem's parameters.

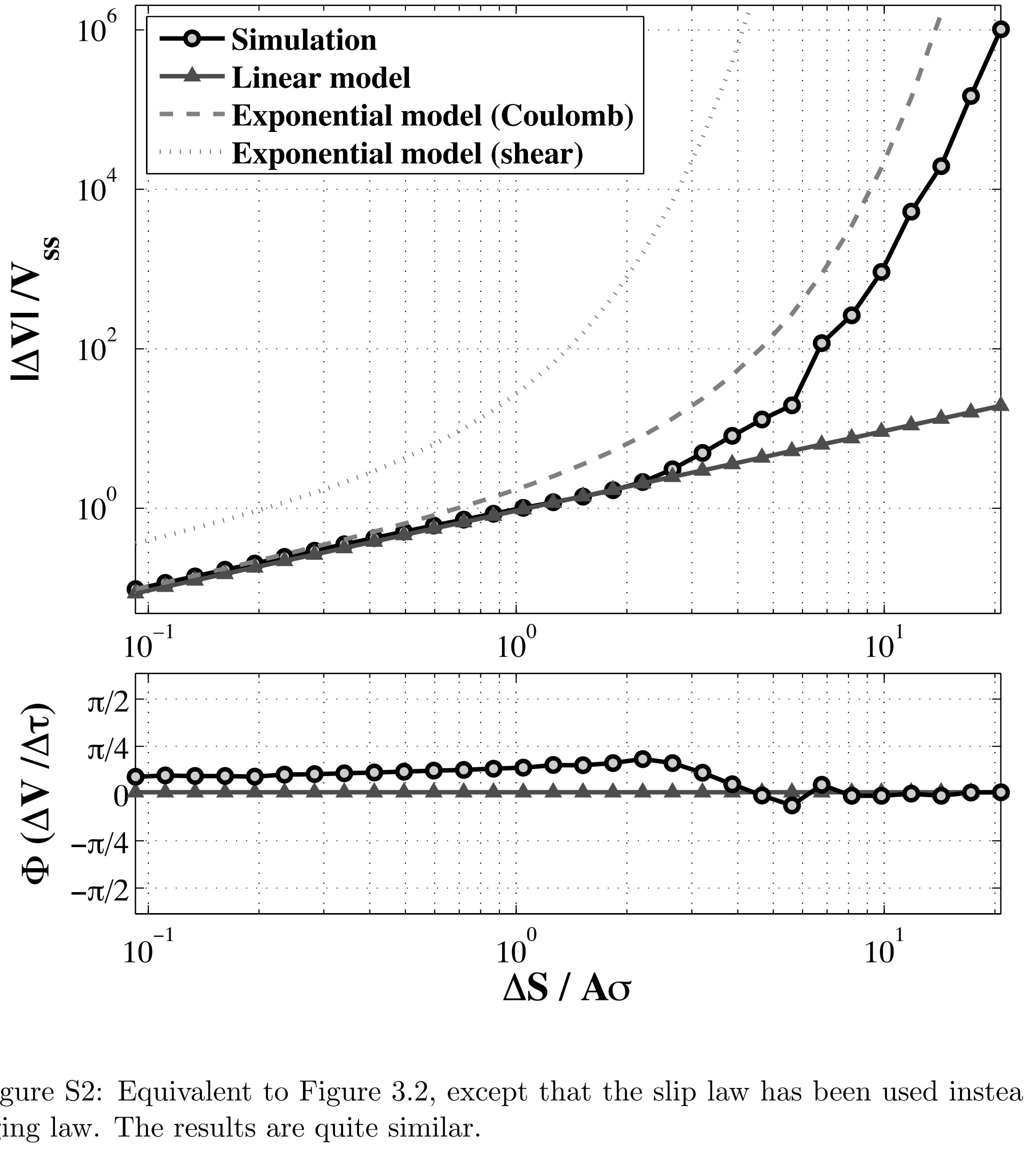


Figure S2: Equivalent to Figure 3.2, except that the slip law has been used instead of the aging law. The results are quite similar.

Chapter 4

Response of rate-and-state seismogenic faults to harmonic shear-stress perturbations

In preparation for the Journal of Geophysical Research - Solid Earth

Thomas Ader,[1,2] Nadia Lapusta,[1] Jean-Philippe Avouac,[1] Jean-Paul Ampuero[1]

[1]Department of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA 91125, USA.

[2]Laboratoire de Géologie, Ecole Normale Supérieure, CNRS, 24 rue Lhomond, 75004 Paris, France.

Abstract

We study the response of the seismicity produced by a 2D seismogenic fault obeying rate-and-state friction laws to harmonically-varying stress perturbations of different periods. Using continuum models of rate-and-state faults made of a seismogenic patch surrounded by creeping areas, we conduct fully dynamic simulations of earthquake sequences using the Boundary Integral CYCLe of Earthquakes approach. We show that when the fault is subjected to harmonic stress perturbations, the corresponding changes in seismicity rate have an amplitude larger than what is predicted by the Coulomb failure model and models based on 1D spring-and-slider systems obeying rate-and-state friction laws. The same conclusions can be drawn for the amplitude of the response of the seismicity to a step-like perturbation of stress. We point out that when inferring fault properties from the response of natural seismicity to stress perturbations, both Coulomb failure and rate-and-state spring-slider models systematically under-estimate the product $a\sigma$, where a is the rate-and-state constitutive parameter relating changes in slip rate to frictional strength, and σ is the effective normal stress on the fault. We suggest that the high sensitivity of the finite fault to external perturbations is due to the sensitivity of the growth of the nucleation zone at the onset of nucleation. The response to harmonic stress perturbations depends on whether the period T of the harmonic stress perturbation is greater or smaller than a characteristic period T_a, similarly to predictions of rate-and-state spring-slider models. At periods $T \gg T_a$, the correlation between the stress perturbation and the seismicity rate is consistent with the Coulomb failure model, i.e., the seismicity-rate variations are proportional to the stress-rate perturbations. At periods $T \ll T_a$, the stress-rate variations are in phase with the stress perturbation, although a gradual phase shift appears as T increases towards T_a. More importantly, as T increases towards T_a, the amplitude of the seismicity-rate variations increases. This would explain observations of variations of seismicity in Nepal, where changes in earthquake frequency correspond with the annually occurring monsoon, whereas there is no such correlation with the semidiurnal Earth tides, although both phenomena yield stress variations of comparable amplitudes. Such a period-dependent behavior of the seismicity

has also been observed in lab experiments. Based on theoretical considerations and analogies with rate-and-state spring-slider models, we propose a semi-analytical expression for the characteristic period T_a, suggesting that T_a is proportional to the ratio D_c/V_{pl}, where D_c is the characteristic slip for state evolution and V_{pl} is the secular loading velocity, and that T_a is independent of the effective normal stress. Comparing the responses of the seismicity to step-like and harmonic perturbations of stress, we highlight the inherent nonlinearity of earthquake generation processes, and therefore the challenges standing in the way of the establishment of an analytical framework capturing the full behavior of the fault based on the physical parameters of the problem.

4.1 Introduction

How a seismogenic fault responds to an applied stress history remains a fundamental question, for which a definitive answer has yet to be set. One of the main obstacles standing in the way toward a solution is the limited range of configurations of stress variations and the resulting seismicity rate arising in nature, which can be constrained from observations. Figure 4.1 depicts this challenge: the responses of seismicity to either a constantly increasing stress, a stress step (e.g., *Gross and Kisslinger*, 1997; *Gross and Bürgmann*, 1998; *Toda et al.*, 1998, 2012), or a periodically varying stress (e.g., *Heki*, 2003; *Cochran et al.*, 2004; *Christiansen et al.*, 2007; *Bollinger et al.*, 2007; *Bettinelli et al.*, 2008; *Ader and Avouac*, 2013) are the most common configurations available. Fortunately, these three configurations actually constitute the standard approach to establish and characterize the transfer function of a linear system, which is fully determined by its ramp, step, and harmonic responses. Seismogenic faults are nonlinear systems since friction is inherently nonlinear, but a lot about earthquake physics can be inferred from their response to these stress configurations. Other mechanisms of earthquake triggering have been mentioned, such as dynamic triggering (e.g., *Hill*, 1993; *Gomberg et al.*, 2003; *Felzer and Brodsky*, 2006), pore fluid motion and induced variations in fault strength (e.g., *Nur and Booker*, 1972; *Bosl and Nur*, 2002) or fluids intrusions (e.g., *Hainzl and Fischer*, 2002; *Cappa et al.*, 2009; *Dahm et al.*, 2010), but are beyond the scope of the present study.

The case of a simple constant loading rate on a fault is probably the most commonly arising in nature. At plate interfaces, for instance, where faults are steadily loaded by the slow motion of tectonic plates, over a time period much longer than the characteristic return period of events of a given magnitude, the stressing rate can be regarded as constant through time and results in an approximately constant seismicity rate, often called “background” rate.

Aftershock sequences following large events are the most frequent type of deviation from this constant background seismicity. Under the premise that aftershocks are statically triggered by the sudden stress change caused by a mainshock, their evolution can be seen as the response of the seismicity to a step-like function in stress. This evolution is characterized by a sudden jump of the seismicity rate immediately after the mainshock, followed by a gradual decay of the seismicity rate with time back to its pre-mainshock level, according to the Omori law (see *Utsu et al.* (2005) for a recent review). The time evolution, amplitude and other characteristics of aftershock sequences have thus been broadly studied, in order to

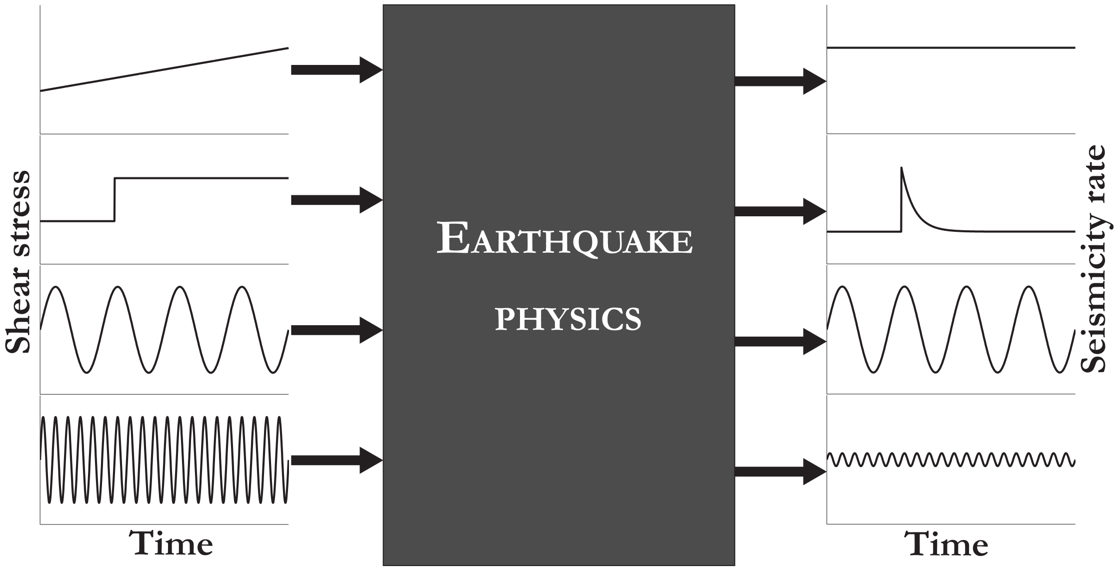


Figure 4.1: Cartoon presenting the challenge underlying today's knowledge of earthquake physics. Earthquake physics would enable one to predict how a fault would react to an imposed stress history, and remains somewhat of a grey box. Few configurations arise naturally and some have been tested in the lab, although lab experiments may not capture the full behavior of natural faults. Most natural faults are loaded at near-constant stress rates which results in constant seismicity rate. A population of faults undergoing a stress step (due to a mainshock) will see its seismicity rate suddenly increase and gradually decay back to the initial seismicity rate, following the Omori law. Seismicity in Nepal gives us the response of seismicity to periodic stresses of comparable amplitudes but different periods, and suggests a larger seismic response to a larger perturbing period, something that cannot be explained by the current seismicity-rate models.

find a way measure fault properties (e.g., *Gross and Kisslinger*, 1997; *Gross and Bürgmann*, 1998; *Toda et al.*, 1998, 2012).

Another case that arises in nature is that of a periodic loading, generally either due to tides (e.g., *Cochran et al.*, 2004), or seasonal loading due to the local hydrological cycle (e.g., *Heki*, 2003; *Christiansen et al.*, 2005, 2007; *Bollinger et al.*, 2007; *Bettinelli et al.*, 2008). Correlation of seismicity with tides has been reported by *Cochran et al.* (2004), who showed a correlation between the occurrence of shallow thrust earthquakes and the occurrence of the strongest tides for global $M > 5.5$ events from the Harvard Centroid Moment Tensor (CMT) catalog. Seasonal variations of seismicity following seasonal variations of stress loading on a fault have been reported at different locations. *Heki* (2003) pointed out possible variations of seismicity in Japan, where the annual variations of snow load seem able to generate annual variations of seismicity. Snow unloading coupled to groundwater recharge was also reported to induce seasonal variations of seismicity in western US volcanic centers (*Christiansen et al.*, 2005), by modifying the stress on the fault by about 5 kPa. The hydrological cycle and associated variations of water load at the surface induce stress variations at depth, which have been reported to induce variations of the seismicity rate on various seismogenic faults. Along the San Andreas fault, *Christiansen et al.* (2007) observed that hydrologically induced stress perturbations of ~ 2 kPa might be sufficient to affect seismicity. Variations of seismicity in the Nepal Himalaya have also been examined, and hydrologically induced stress variations on the Main Himalayan Thrust fault (MHT) of amplitude of about 3 kPa appear to produce seasonal variations of the seismicity rate of amplitude $\sim 40\%$ (*Ader and Avouac*, 2013), in phase with the variations of stress rate on the fault (*Bollinger et al.*, 2007; *Bettinelli et al.*, 2008), while the seismicity rate appears to remain unresponsive to tidal stress variations yet of similar amplitude (*Bettinelli et al.*, 2008; *Ader and Avouac*, 2013). This would indicate a period-dependent response of the Nepalese seismicity, with less sensitivity at tidal periods than to the annually occurring monsoon.

Such a period-dependent response of faults has been observed in various laboratory experiments, where sample faults undergo harmonic stress variations (*Lockner and Beeler*, 1999; *Beeler and Lockner*, 2003; *Savage and Marone*, 2007, 2008). All of these studies observe two distinct regimes of response of the fault depending on the perturbation's period. At periods larger than a critical period, the seismicity rate on the sample fault appears to be directly proportional to the rate of harmonically varying stress. In such a regime, the amplitude of the seismicity-rate variations are inversely proportional to the perturbation's period. At shorter periods, sample faults appear to have a different behavior. Lab experiments by *Lockner and Beeler* (1999) and *Beeler and Lockner* (2003) suggested a slightly period-dependent response, with the correlation between the timing of events and the stress perturbation increasing with the perturbation's period. This period-dependent response at shorter periods could explain the observations in Nepal, whereas changes in earthquake frequency correspond with the annually occurring monsoon, there is no such correlation with Earth tides, which oscillate back-and-forth twice a day.

The multitude of various efforts deployed have contributed to our increasing understanding of earthquake physics. The grey box of earthquake mechanics depicted in Figure 4.1 becomes more transparent as new discoveries are made, bringing to light the underlying gearwheel. The mechanisms seemingly acting on faults are well described in the formalism of rate-and-state friction (*Dieterich*, 1978, 1979a,b; *Ruina*, 1983), where the friction between two rock surfaces of within a granular rock layer depends on the relative slip velocity and a state variable evolving with time. The rate-and-state friction laws were established in order to reproduce the observations that the onset of frictional sliding in lab experiments is a time-dependent process, and introduce a time-dependent failure mechanism for the generation of earthquakes.

Using this formalism and modeling the earthquake generation process on faults with a simple 1D spring-and-slider system obeying rate-and-state friction laws, *Dieterich* (1994) proposed an analytical expression linking the stress history on a fault to the expected resulting seismicity rate. To some extent, some of the observations described earlier can be explained by the predictions of this Spring-slider Rate-and-state Model (SRM). The Omori law for the decay of aftershocks rate with time is well reproduced by the SRM (*Dieterich*, 1994), although it requires a near-lithostatic pore pressure where aftershocks nucleate in order to quantitatively explain the typical duration of aftershock sequences (e.g., *Gross and Kisslinger*, 1997; *Gross and Bürgmann*, 1998; *Toda et al.*, 1998, 2012). In the case of harmonic stress perturbations, the SRM explains the phase shift between the stress perturbations and the resulting seismicity-rate variations (e.g., *Beeler and Lockner*, 2003). Applied to the observations in Nepal, the SRM also requires a near-lithostatic pore pressure in the seismogenic zone in order to explain the amplitude of the correlation between the seismicity rate and the monsoon-induced variations of stress (*Bettinelli et al.*, 2008).

As has been highlighted in the case of non-volcanic tremors (*Ader et al.*, 2012b), near-lithostatic pore pressures require specific fault properties for the nucleation sizes to remain consistent with the occurrence of the smallest earthquakes recorded. Besides, near-lithostatic pore pressure and corresponding effective normal stresses are orders of magnitude below the values reported from afterslip studies in various tectonic contexts (*Hearn et al.*, 2002; *Miyazaki et al.*, 2004; *Perfettini and Avouac*, 2004, 2007; *Hsu et al.*, 2006, 2009a,b; *Fukuda et al.*, 2009; *Barbot et al.*, 2009). Moreover, the SRM does not explain the period-dependent response of the seismicity to harmonic stress perturbations at shorter periods as was observed in lab experiments (*Lockner and Beeler*, 1999; *Beeler and Lockner*, 2003) and in Nepal (*Bettinelli et al.*, 2008; *Ader and Avouac*, 2013).

In this study, we therefore drop the spring-slider approximation and study the response of a 2D seismogenic fault obeying rate-and-state friction laws to harmonically-varying stress perturbations of different periods. In other words, we investigate the harmonic response of a 2D seismogenic finite fault. To do so, we conduct fully dynamic simulations of earthquake sequences (*Lapusta et al.*, 2000; *Lapusta and Rice*, 2003; *Lapusta and Liu*, 2009; *Noda and Lapusta*, 2010) on a seismogenic patch of finite size surrounded by creeping areas, and undergoing stress perturbation. Our goal is to compute the corresponding changes of the seismicity rate. The motivation for studying the response of a finite fault comes from the study of *Kaneko and Lapusta* (2008) that showed that finite faults have different reposes to shear stress than spring-slider models. The present study mostly focuses on the response of such a finite fault to harmonic perturbations, but we also present some results of the step response of the finite fault.

In the following, we start by briefly reviewing in sections 4.2 and 4.3 the main principles of the two most common models relating the stress history to the expected seismicity rate on a fault — the Coulomb Failure Model (CFM) and the SRM — and present their respective predictions for the response of seismicity to a step-like and a harmonic stress perturbation. We then present the finite-fault simulations that constitute the core of this study by first describing the methodology in section 4.4 and then the frequency response of finite faults in section 4.5. Section 4.6 then introduces results of the step response highlighting properties of the finite fault revealed by the frequency response, as well as illustrating the general nonlinearity of the response. We propose an interpretation of the results in section 4.7 and conclude in section 4.8.

4.2 Response of seismicity in the Coulomb Failure model

The CFM is probably the simplest way to conceptualize the relation between stressing and earthquake occurrence on a fault. It assumes that whenever the Coulomb stress $S = \tau - \mu\sigma_{\text{eff}}$, where μ is the friction coefficient, τ the shear stress and σ_{eff} the effective normal stress (i.e., the normal stress reduced by the pore pressure), reaches a threshold value called the Coulomb Failure Stress (CFS), the fault produces an earthquake and the stress on the fault drops to a lower value. Assuming in addition a population of faults on which the pre-stresses are uniformly distributed up to the CFS, the observed seismicity rate $R(t)$ is therefore proportional to the Coulomb stress rate $\dot{S}(t)$, and thus to the shear-stress rate $\dot{\tau}(t)$ when the normal stress is kept constant. Since the seismicity rate cannot have negative values, this relation of direct proportionality remains true only as long as the Coulomb stress keeps increasing. If it starts decreasing, there will not be any earthquakes until it grows back to a value equal to its last maximum. Denoting by $S_f(t)$ the increasing envelope of the Coulomb stress $S(t)$ (see Figure 4.A.1a in appendix 4.A for a representation of the equivalent functions $\tau(t)$ and $\tau_f(t)$), the seismicity rate $R(t)$ in the CFM can be simply written as:

$$R(t) \propto \dot{S}_f(t). \tag{4.1}$$

In the case of a step-like change of stress on the fault, this model implies a simple impulse change of the seismicity rate, and therefore does not reproduce the Omori law for the decay of aftershocks with time.

The case of the response to a harmonic stress perturbation is more insightful. The seismicity rate in the CFM is proportional to the stress rate as has been reported for the seasonal variations of seismicity in Nepal (*Bettinelli et al.*, 2008), and for the results of laboratory experiments at larger perturbing periods (*Lockner and Beeler*, 1999; *Beeler and Lockner*, 2003; *Savage and Marone*, 2007, 2008). Given a background loading rate of shear stress $\dot{\tau_a}$, superimposing harmonic variation of shear stress of amplitude $\Delta\tau$ might cause the resulting shear stress to periodically decrease, if the period T of the perturbation is short enough (Figure 4.A.1a). We show in appendix 4.A that, depending whether the period T of the perturbation is greater or smaller than the critical period $T_\tau = 2\pi\Delta\tau/\dot{\tau_a}$, the relative amplitude of the variations of seismicity rate is

$$\frac{\Delta R}{r} = \frac{T_\tau}{T} \quad \text{when} \quad T \geq T_\tau, \tag{4.2}$$

and

$$\frac{\Delta R}{r} = 2\sqrt{\pi}\sqrt{\frac{T_\tau}{T}} \quad \text{when} \quad T \ll T_\tau. \tag{4.3}$$

Equations (4.2) and (4.3) indicate that at all periods, the amplitude of the seismicity response increases as the period decreases, in contradiction with observations. In the case of

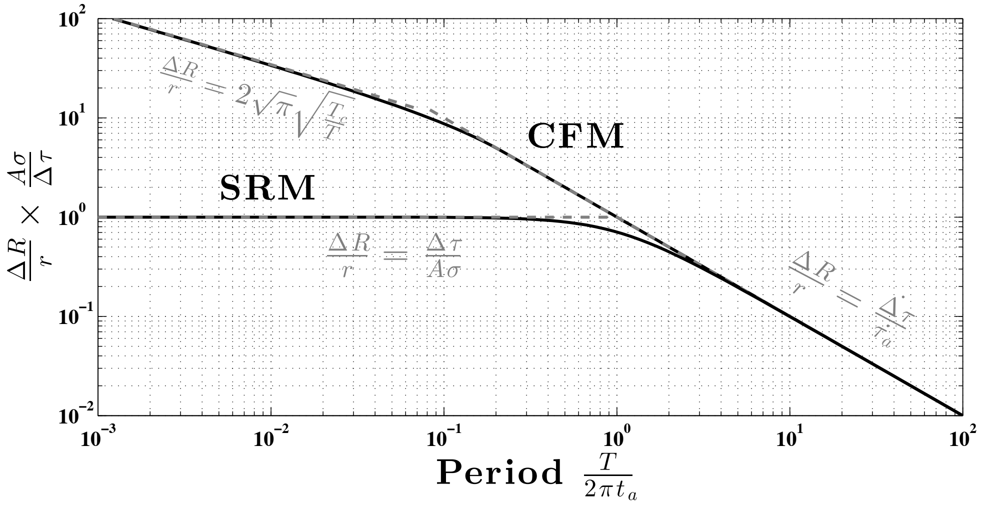


Figure 4.2: The amplitude of seismicity rate variations for different periods of harmonic variations in shear stress, according to the CFM and SRM. The black curves represent the exact solutions while the grey dashed lines represent the asymptotic behavior, with corresponding equations indicated on the plot, and derived in the appendices.

Nepal, for instance, this would imply a greater response of the seismicity to the semidiurnal tidal loading than to the annual forcing. This is also incompatible with the results of lab experiments at shorter perturbing periods by *Lockner and Beeler* (1999) and *Beeler and Lockner* (2003). The predicted amplitude of the variations of seismicity rate $\Delta R/r$ in response to a harmonic perturbation of shear stress of amplitude $\Delta\tau$ by the CFM is plotted in Figure 4.2 as a function of the perturbing period T.

4.3 Response of seismicity in the SRM

The major issue with the CFM is that it does not contain any time-dependent mechanism for the earthquake nucleation process that could reproduce the gradual decay of aftershocks rate following a mainshock. *Dieterich* (1994) proposed an alternative model of the seismicity rate on a fault based on a 1D spring-and-slider system following rate-and-state friction laws.

In the rate-and-state formalism, the evolution of the friction coefficient μ between two rock surfaces or gouge layers, or below the slider in the case of the SRM, logarithmically depends on the slip rate V and a state variable θ (*Dieterich*, 1978, 1979a,b; *Ruina*, 1983):

$$\mu = \mu^* + a \ln \frac{V}{V^*} + b \ln \frac{\theta V^*}{D_c}, \tag{4.4}$$

where μ^* is the reference friction coefficient corresponding to the reference slip velocity V^*,

D_c is the characteristic slip for state evolution (e.g., *Dieterich*, 1978, 1979a,b; *Ruina*, 1983; *Rice and Ruina*, 1983; *Dieterich and Kilgore*, 1994), and $a > 0$ and $b > 0$ are rate-and-state constitutive fault parameters with $a - b < 0$ so that the system has a rate-weakening rheology. The state variable θ can be interpreted as the average age of the population of contacts between two surfaces and evolves according to the "aging law" (e.g., *Marone*, 1998):

$$\frac{d\theta}{dt} = 1 - \frac{V\theta}{D_c}. \tag{4.5}$$

Note that the state variable evolves in time even if there is no relative motion between rocks in contact.

In order to derive a relation between the seismicity rate on a population of faults undergoing a time-varying stress, the SRM of *Dieterich* (1994) makes a few assumptions. The model assumes that the times to failure are uniformly distributed, and that variations of stress on the faults simply modify the time to failure. Besides, it is assumed that, at the onset of rupture, the velocity on the fault is large enough so that $V\theta/D_c \gg 1$, thus reducing equation (4.5) to $d\theta/dt = -V\theta/D_c$. The relations obtained are reviewed in appendix 4.B for the case of a constant normal stress. For a stress step mimicking a sudden stress change produced by a nearby mainshock, SRM's relations successfully reproduce the Omori law for the time decay of aftershocks (equation (12) from *Dieterich* (1994)). The cumulative number of events following a stress step a time $t = 0$ is given by:

$$N(t) = rt + rt_a \ln\left[e^{\Delta\tau/a\sigma} + \left(1 - e^{\Delta\tau/a\sigma}\right)e^{-t/t_a}\right]\mathcal{H}(t), \tag{4.6}$$

where

$$t_a = a\sigma/\dot{\tau}_a \tag{4.7}$$

is the characteristic relaxation time of the seismicity rate following a stress step (i.e., the characteristic duration of an aftershock sequence), and $\mathcal{H}(t)$ is the Heavyside function, i.e., $\mathcal{H}(t) = 0$ for $t < 0$ and $\mathcal{H}(t) = 1$ for $t \geq 0$. This model is able to link the parameters of the rate-and-state formalism and the empirical Omori's law for the time decay of aftershocks. It has therefore fostered numerous studies inferring fault properties from observations of aftershock sequence decays (*Gross and Kisslinger*, 1997; *Gross and Bürgmann*, 1998; *Toda et al.*, 1998, 2012, e.g.,). With the right estimates of the stress step amplitude $\Delta\tau$ and the secular stress rate $\dot{\tau}_a$ on the fault, a fit of equation (4.6) to the observed cumulative number of events in the aftershock sequence yields estimates of the product $a\sigma$. The fault parameter a has been measured in laboratory experiments (*Dieterich* (1994) found 0.005 to 0.012) and, assuming that these laboratory-derived values can be applied to real faults, $a\sigma$ leads to estimates of the effective normal stress at the depth of the aftershock sequence.

This exercise has been performed on numerous aftershock sequences, systematically leading to effective normal stresses one to several orders of magnitude below the lithostatic pressure at studied depths. Using the temporal evolution of the 1992 Landers event aftershocks for instance, *Gross and Kisslinger* (1997) constrained $a\sigma$ between 23 and 47 kPa. Following the same procedure on the 1989 Loma Prieta earthquake, *Gross and Bürgmann* (1998) estimated $a\sigma = 11$ to 330 kPa, based on the value of t_a in the fit. In Japan, *Toda et al.* (1998) computed $a\sigma = 35$ kPa for the time decay of the aftershocks of the 1995 Kobe event, and *Toda et al.* (2012) reported $a\sigma = 10$ kPa for the Joshua Tree and Landers earthquakes. Noting that an effective normal stress equal to the hundreds of MPa of overburden pressure at seismogenic depth would lead to values of a much less than the range of laboratory derived values (*Dieterich*, 1994), these studies thus concluded that the prevailing effective normal stress should be of the order of a few MPa, i.e., two orders of magnitude below the lithostatic value. To explain this difference, they appealed to near-lithostatic pore pressures at seismogenic depths, which would reduce the effective normal stress by a few orders of magnitude.

Regarding the response of seismicity to harmonic stress perturbations, one can conceptually understand how the time-dependent failure mechanism introduced by the rate-and-state laws in the SRM might be able to dampen the frequency response at shorter periods: if the nucleation time of events t_a is much larger than the stress perturbation period T, the stress variations seen by the seismicity is smoothed out, and the amplitude of the seismicity response is reduced. Conversely, in the opposite case where nucleation time t_a is much smaller than the stress perturbation period T, the existence of a nucleation time will simply introduce a phase shift $\Phi \sim 2\pi t_a/T$ between the stress rate and the seismicity rate, rapidly negligible with increasing T, but should not produce a response much different from the one predicted by the CFM. The complete derivation of the harmonic response for the SRM is detailed in appendix 4.B. As has been noticed in previous studies (e.g., *Lockner and Beeler*, 1999; *Beeler and Lockner*, 2003) and described before, this model indeed predicts two different behaviors of the seismicity response (Figure 4.2), depending on whether the period of the perturbation T is larger or smaller than the characteristic period T_a, defined as

$$T_a = 2\pi t_a = 2\pi \frac{a\sigma}{\dot{\tau}_a}. \tag{4.8}$$

As expected, the critical period T_a is directly related to the characteristic relaxation time t_a.

For perturbation periods T much larger than T_a, the seismicity responds in a Coulomb Failure fashion. The seismicity rate is directly proportional to the stress rate as long as it remains positive, otherwise, a seismicity quiescence is observed, exactly following the predictions of the CFM (see equations (4.B.14) and (4.B.15) in the appendix). In the case of perturbations of period $T \ll T_a$, as expected, the response of the seismicity is lower than the predictions of the CFM. The amplitude of the seismicity-rate variations actually becomes independent of the period, and the seismicity rate in response to a shear stress perturbation $\tau(t) = \dot{\tau_a} t + \Delta\tau \sin \omega t$ can be written as (equation (4.B.13) in the appendix):

$$R(t) \propto \exp\left(\frac{\Delta\tau}{a\sigma} \sin \omega t\right). \tag{4.9}$$

In the case where the amplitude of the perturbations is such that $\Delta\tau \ll a\sigma$, the relative amplitude of the seismicity rate around its unperturbed value is simply $\Delta R/r = \Delta\tau/a\sigma$ (equation (4.B.7) in the appendix). This small-perturbation configuration is the one considered later in the finite-fault simulations, in order to avoid introducing any extra source of nonlinearity. Like for aftershock-rate predictions, these simple expressions have enabled studies to determine values of the product $a\sigma$ from real cases of variations of seismicity. Looking at triggering of earthquakes by tides worldwide, *Cochran et al.* (2004) fitted the amplitude of induced variations of seismicity with the SRM and provided a range of values for $a\sigma$ between 48 and 110 kPa, with a best fit of 64 kPa, values of the same order of magnitude as the ones inferred from the study of aftershock sequences. Based on the results of the SRM, *Bettinelli et al.* (2008) reported that for the Nepalese seismicity to be able to respond to seasonal variations of surface water load, extremely low values of $a\sigma$ (between 3 and 8 kPa) were required. These results would thus also suggest extremely low effective normal stresses at the seismogenic depth of the faults, at least at places where aftershocks nucleate.

The SRM reproduces, at least qualitatively, the usually observed time decay of aftershocks rates, but an issue remains for the frequency response. Even though the response of the seismicity at short periods is dampened compared to what the CFM would predict, the response still remains at least as large when the period gets shorter. Therefore, the SRM cannot explain the observations in Nepal (*Bollinger et al.*, 2007; *Bettinelli et al.*, 2008; *Ader and Avouac*, 2013) and the results of lab experiments by *Lockner and Beeler* (1999) and

Beeler and Lockner (2003). Besides, it seems to require a near-lithostatic pore pressure, i.e., extremely low normal stresses at seismogenic depths in order to quantitatively explain both aftershock sequences and response to periodic perturbations. As mentioned earlier, low effective normal stresses might be a problem for the nucleation of the smallest recorded events, and they are not observed in afterslip studies.

We thus investigate whether the rate-and-state law alone is able to reproduce the results observed for the step and harmonic responses of seismicity, by setting aside the spring-slider approximation, and examining what would be the behavior of a 2D fault with rate-and-state friction under such stress perturbations.

4.4 Modeling a finite rate-and-state fault

We study the behavior of the 2D fault schematically represented in Figure 4.3, infinite in one direction, and consisting of a potentially seismogenic rate-weakening patch($a - b < 0$), surrounded by rate-strengthening areas ($a - b > 0$). The fault is 3 km long and, unless stated otherwise, discretized into 6000 cells of 0.5 m each. The evolution of the slip on this finite rate-and-state fault is simulated using the Boundary Integral CYCLe of Earthquakes (BICYCLE) approach (*Lapusta et al.*, 2000; *Lapusta and Rice*, 2003; *Lapusta and Liu*, 2009; *Noda and Lapusta*, 2010). Several thousands of earthquakes are simulated in order to have a statistically significant number of events. The fault is loaded at constant slip velocity $V_{pl} = 1$ cm/yr (unless indicated otherwise) on both sides, and the rate-and-state parameters are $a = 0.008$ on the entire fault, $b = 0.012$ on the seismogenic patch and 0.004 in the creeping zone, $D_c = 5$ μm, and the reference friction coefficient is $\mu^* = 0.6$ at slip rate $V^* = 10^{-6}$ m/s. The medium has a shear modulus of $G = 30$ GPa and, unless noted otherwise, $\sigma = 5$ MPa. In fully dynamic simulations of 2-D antiplane earthquakes sequences, resolving the cohesive zone size Λ_0 with 3 to 5 spatial cells is the more stringent requirement for the aging formulation of rate-and-state friction and typical rate-and-state parameters (*Lapusta and Liu*, 2009). For a fault interface governed by rate-and-state friction laws, Λ_0 can be expressed as (*Palmer and Rice*, 1973; *Day et al.*, 2005; *Lapusta and Liu*, 2009)

$$\Lambda_0 = \frac{GD_c}{b\sigma}. \tag{4.10}$$

In our case, $\Lambda_0 = 2.5$ m, which justifies our choice of 0.5 m for the cell size in our simulations. Unless indicated otherwise, the seismogenic patch is 500 m long. Our choice of $\sigma = 5$ MPa is motivated by the small stress perturbations we would like to study and the fact that the

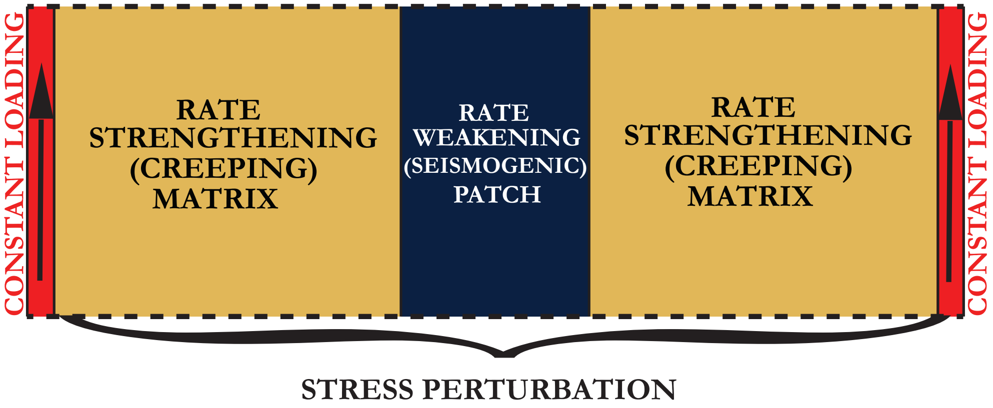


Figure 4.3: Schematics of the finite fault used in the numerical simulations. The fault is infinite in one direction, and consists of one rate-weakening (seismogenic) patch embedded within a rate-strengthening medium. The whole fault is loaded on both sides at a constant slip velocity. The stress perturbation (either a step function or a harmonic perturbation) is applied over the entire fault. Unless otherwise indicated, the fault has a length of 3 km, and the seismogenic patch at the center is 500 m long. The normal stress is held constant at $\sigma = 5$ MPa. The rate-and-state fault parameters are: $a = 0.008$ over the entire fault, $b = 0.012$ in the seismogenic patch (so that $a-b < 0$) and $b = 0.004$ in the rate-strengthening region (so that $a - b > 0$), $D_c = 5$ μm, and the reference friction coefficient is $\mu^* = 0.6$ at the reference slip velocity $V^* = 10^{-6}$ m/s.

SRM predicts an amplitude not greater than $\Delta\tau/a\sigma$ for the response of the seismicity rate to a stress perturbation of amplitude $\Delta\tau$. In order to look at stress perturbations of the order of 3 kPa, as has been estimated for the monsoon-induced stresses in Nepal, we settle for the largest normal stress able to bring out a response of the seismicity large enough to be detected.

Figure 4.4a shows the natural evolution of slip along the fault over about 2 years, without any exterior stress perturbation. In order to represent both the interseismic and coseismic slip in Figure 4.4a, the slip is plotted every 0.01 years if the fault is in the interseismic regime and every 0.02 seconds if it is in the coseismic regime. We consider that a seismic event is occurring on the fault when the maximum velocity on the fault is greater than 1 cm/s, many orders of magnitude greater than the loading velocity of 1 cm/yr. The linear magnitude M_{lin} of each of the seismic events on the fault is indicated in Figure 4.4a, and it is defined as:

$$M_{\text{lin}} = \frac{2}{3}\log_{10}\mathcal{M}_{\text{lin}} - 6.7, \tag{4.11}$$

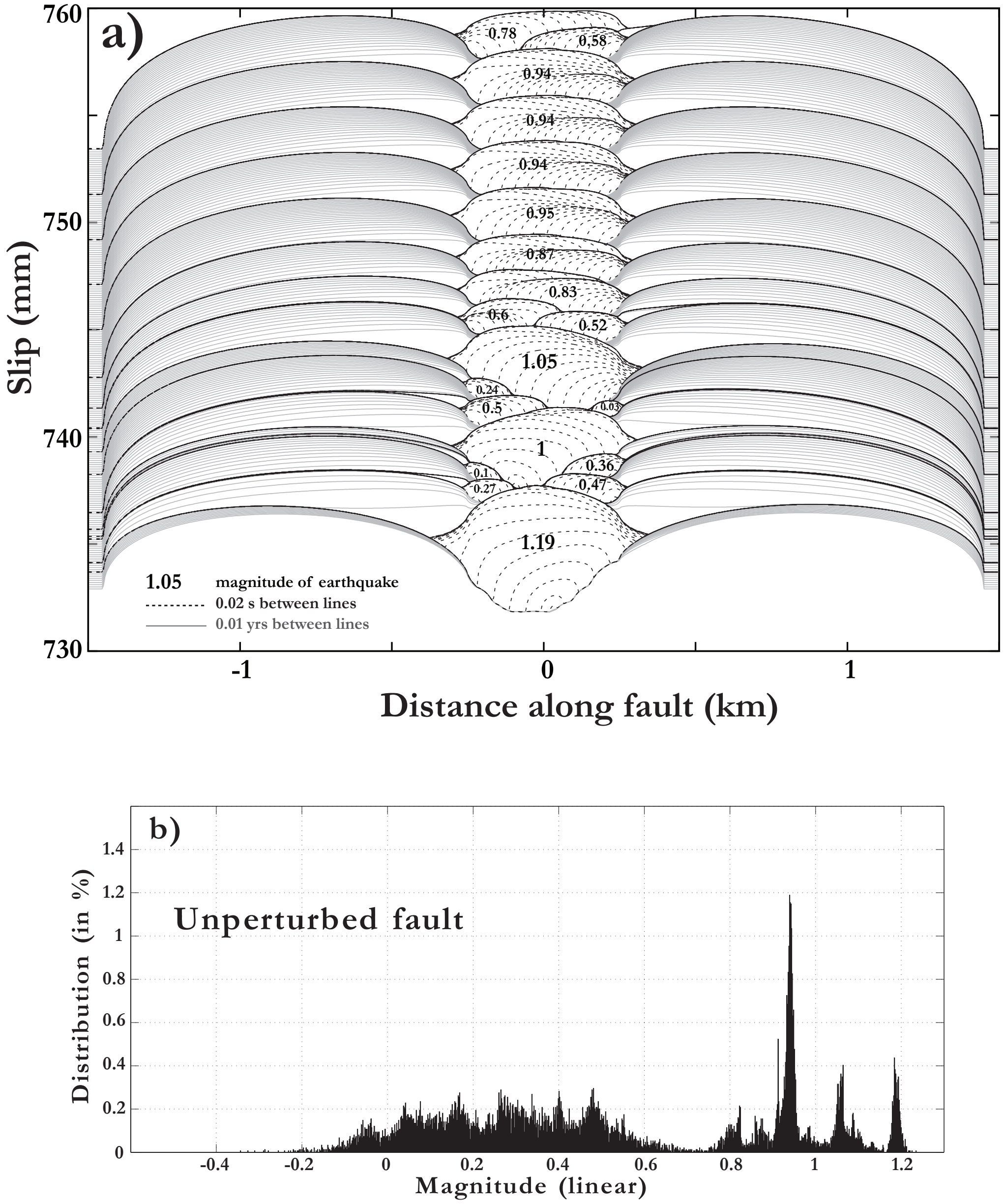


Figure 4.4: a) Evolution of slip on the unperturbed fault with time. The linear magnitude of the seismic events produced is indicated on each event. If the fault is experiencing a seismic event, the slip on the fault is plotted every 0.02 s (dashed black lines), while during the interseismic period, slip is only plotted every 0.01 yrs (plain grey lines). b) Magnitude distribution of events produced by the fault. The earthquake catalog contains a total of about 15,000 events, covering about 1700 years of evolution of the fault.

where the linear moment $\mathcal{M}_{\mathrm{lin}}$ of a seismic event on this fault is

$$\mathcal{M}_{\mathrm{lin}} = G \int_{fault} s(x)\, dx, \tag{4.12}$$

$x \in [-1500; 1500]$m is the position along the fault and $s(x)$ is the slip on the fault at position x. Note that initial conditions assumed on the fault affect only several first events; afterwards, the fault behavior becomes independent of initial conditions (e.g., *Lapusta and Liu*, 2009). In Figure 4.4a and other similar Figures, the first events shown occur after at least 100 other events that have been removed.

When evolving only under constant background loading, the fault can produce a fairly broad complexity of event sizes and earthquake patterns (Figure 4.4). Some events rupture the entire seismogenic patch at once, still producing events of various magnitudes, while some smaller events only rupture an edge of the patch. The complexity produced by this fault is due to the fact that the critical size necessary to nucleate seismic events on the seismogenic patch is much smaller than the total length of the seismogenic patch. The critical nucleation size has been analyzed in numerous studies (e.g., *Ruina*, 1983; *Dieterich*, 1992; *Rubin and Ampuero*, 2005), and can be written as:

$$h^* = \frac{GD_c}{F(a,b)\sigma}, \tag{4.13}$$

where the function $F(a,b)$ of the fault parameters a and b refers to different models estimating the critical nucleation size, and can be $F(a,b) = a - b$ (*Ruina*, 1983), $F(a,b) = b$ (*Dieterich*, 1992), or $F(a,b) = b$ when $a/b < 0.37$ and $F(a,b) = \pi/2 \times (b-a)^2/b$ when $a/b > 0.5$ (*Rubin and Ampuero*, 2005). In our case, $a/b \approx 0.67$, and taking the estimation by *Rubin and Ampuero* (2005) yields

$$h^* = \frac{2}{\pi}\frac{b}{(b-a)^2}\frac{GD_c}{\sigma} \approx 15 \text{ m}. \tag{4.14}$$

In order to illustrate the event complexity, Figure 4.4b represents the distribution of linear magnitudes of all the seismic events produced by this fault. To the first order, this distribution is bimodal: events with $M_{\mathrm{lin}} > 0.75$ rupture the entire seismogenic patch while events with $M_{\mathrm{lin}} < 0.75$ only rupture an edge of the seismogenic patch. As illustrated in Figure 4.4a, the fault exhibits irregularity, a fact supported in Figure 4.4b by the existence of various peaks in the magnitudes distribution at $M_{\mathrm{lin}} > 0.75$. Larger events of $M_{\mathrm{lin}} \sim 1.05$ and $M_{\mathrm{lin}} \sim 1.2$ alternate with smaller foreshocks and aftershocks with $M_{\mathrm{lin}} < 0.75$ and

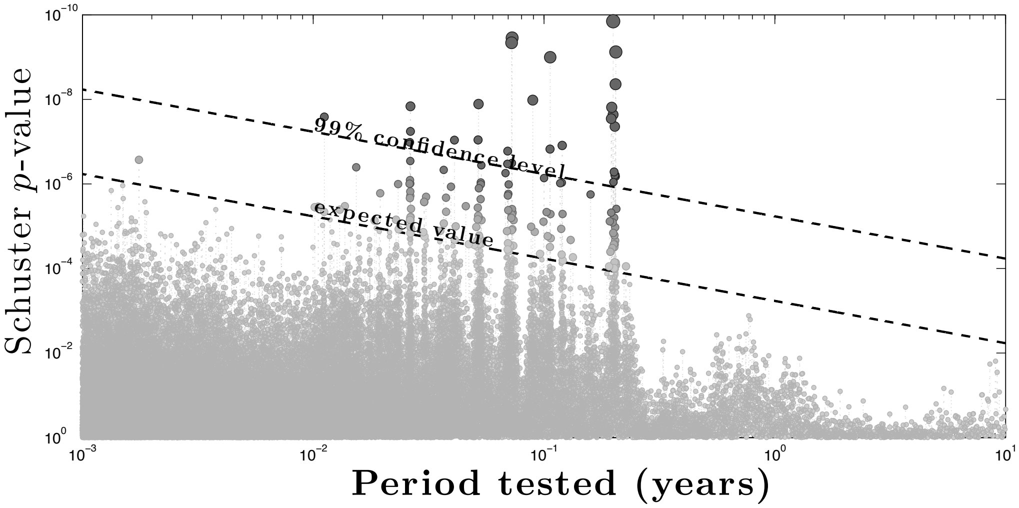


Figure 4.5: Schuster spectrum for the unperturbed fault (Figure 4.4). The spectrum is computed for the timing of about 15,000 events spanning 1700 years of history on the fault. The Schuster p-value indicates the probability that a periodicity is observed by chance in the timing of events. The "expected value" line indicates the expected value of the Schuster p-value, while points above the "99% confidence level" line have a probability above 99% to be real periodicities in the catalog, rather than being observed by chance. The periodicity at $T \approx 0.2$ years corresponds to the return period of $M_{\text{lin}} \approx 0.94$ events, indicating that these events happen quite periodically on the fault, while periodicities at smaller periods are harmonic of this period.

sometimes other small events during the interseismic period (first 12 events in Figure 4.4a), whereas events of $M_{\text{lin}} \sim 0.94$ that also rupture the entire patch seem to almost periodically follow each other without any smaller events in-between (events 13 to 18 in Figure 4.4a).

This is even more obvious when looking at periodicities in the timing of events produced by this fault: Figure 4.5 shows the Schuster spectrum (*Ader and Avouac*, 2013) computed for the timing of the $\sim 15,000$ events from the catalog generated by the simulation. This spectrum shows the periodicities in the timings of events on the fault: peaks above the "99% confidence level" line indicate that a periodicity at the tested period exists in the timing of events in the catalog at more than a 99% level of confidence. The Schuster spectrum thus indicates that this fault has a natural periodicity around period of $T = 0.02$ years. The periodicities at smaller periods are simply harmonics of this period (*Ader and Avouac*, 2013).

To understand the origin of this natural periodicity, one can estimate the return period of events of a given linear moment $\mathcal{M}_{\text{lin}}$, by supposing that such events release all the moment deficit accumulated on the seismogenic patch from the constant loading:

$$T_{\mathrm{RET}}(\mathcal{M}_{\mathrm{lin}}) = \frac{\mathcal{M}_{\mathrm{lin}}}{GWV_{pl}}, \tag{4.15}$$

where $W = 500$ m is the length of the seismogenic patch and $V_{pl} = 1$ cm/yr is the loading plate velocity. In terms of linear moment magnitude M_{lin}, and with the parameters of the simulation, equation (4.15) becomes:

$$T_{\mathrm{RET}}(M_{\mathrm{lin}}) = 10^{1.5M_{\mathrm{lin}}-2.08}. \tag{4.16}$$

Using equation (4.16), the expected return periods of events rupturing the entire seismogenic patch corresponding to the main peaks in Figure 4.4b would be $T_{\mathrm{RET}}(M_{\mathrm{lin}} = 0.94) \approx 0.21$ years, $T_{\mathrm{RET}}(M_{\mathrm{lin}} = 1.05) \approx 0.31$ years and $T_{\mathrm{RET}}(M_{\mathrm{lin}} = 1.2) \approx 0.53$ years. Periods corresponding to the return periods of $M_{\mathrm{lin}} = 1.05$ and $M_{\mathrm{lin}} = 1.2$ events do not appear in the Schuster spectrum of the seismicity, indicating that events of magnitude $M_{\mathrm{lin}} > 1$ happen in a fairly chaotic manner. The strong periodicity appearing at 0.2 years in Figure 4.5 corresponds to the expected return period of $M_{\mathrm{lin}} \sim 0.94$ events, implying that these happen at quite regular time intervals on the fault.

4.5 Response of a rate-and-state seismogenic fault to harmonic shear-stress perturbations

4.5.1 Method

In order to study the response of the fault presented in the previous section to harmonic stress perturbations, we simulate the evolution of the slip on this fault with the exact same fault parameters as previously described, while superimposing a harmonically varying stress over the entire fault (both the rate strengthening and the rate weakening parts):

$$\Delta\tau(t) = \Delta\tau \sin \omega t, \tag{4.17}$$

where the amplitude of the stress perturbation is held constant at $\Delta\tau = 3$ kPa, as inferred for Nepal (*Bettinelli et al.*, 2008). The ratio $\Delta\tau/a\sigma = 0.075 \ll 1$, indicating that we are in the case of small perturbations according to section 4.3. We run simulations for perturbations at periods varying between 10^{-6} and 10 years. In order to estimate the amplitude of the variations of seismicity and their phase, we stack the times of N events from the simulated catalog over the perturbing period. We then compute the seismicity rate over the perturbing period T by dividing the period into $B = 32$ bins of equal duration T/B and counting the number of events falling within each bin. We finally normalize this stacked seismicity rate by its mean, so that the average seismicity rate is equal to 1. Given the number of events N in the catalog and the number of bins B, the variance of the normalized number of events falling within each bin is (e.g., *Ader and Avouac*, 2013):

$$\sigma_B^2 = (B-1)/N. \tag{4.18}$$

For each simulated catalog, we then fit the seismicity rate with a function qualitatively following the predictions of *Dieterich* (1994):

$$\frac{R(t)}{r} = \frac{e^{\beta \sin(\omega t - \Phi)}}{\left\langle e^{\beta \sin \omega t} \right\rangle}, \tag{4.19}$$

where the amplitude β and phase Φ of the response are determined to fit the results of each simulation. Note that $\beta = \Delta\tau / a\sigma$ according to the SRM for periods smaller than T_a; here, we do not impose β but rather determine it from the simulation results. The notation $\langle . \rangle$ refers to the mean of the function.

Figure 4.6 illustrates this process: it shows, for 4 different simulations with different perturbing periods, the stacked seismicity rate (black circles with error bars corresponding to σ_B in equation (4.18)) and the result of the fit. This Figure shows that, although the perturbation is of a small amplitude, the resulting variations of the seismicity rate can actually be fairly large, and that a phase shift appears as the perturbing period increases (recall that the shear stress perturbation is a sine function, and so it is 0 at time equals 0). The amplitude of the seismicity rate variations thus appears to be dependent on the perturbation period in non-monotonic manner, being higher at $T = 0.027$ years than at both $T = 0.0027$ years and $T = 1$ year.

In order to estimate uncertainties on β and Φ, we use analytical estimates for the case of small variations of the seismicity rate, i.e., $\beta \ll 1$. In this case, equation (4.19) can be linearized:

$$\frac{R(t)}{r} = 1 + \beta \sin(\omega t - \Phi) = 1 + m_1 \sin \omega t + m_2 \cos \omega t, \tag{4.20}$$

where m_1 and m_2 are two linear parameters determined by the fit to the seismicity rate, such that $\beta = \sqrt{m_1^2 + m_2^2}$ and $\Phi = \tan^{-1}(m_2/m_1)$. With parameters m_1 and m_2, the fit is

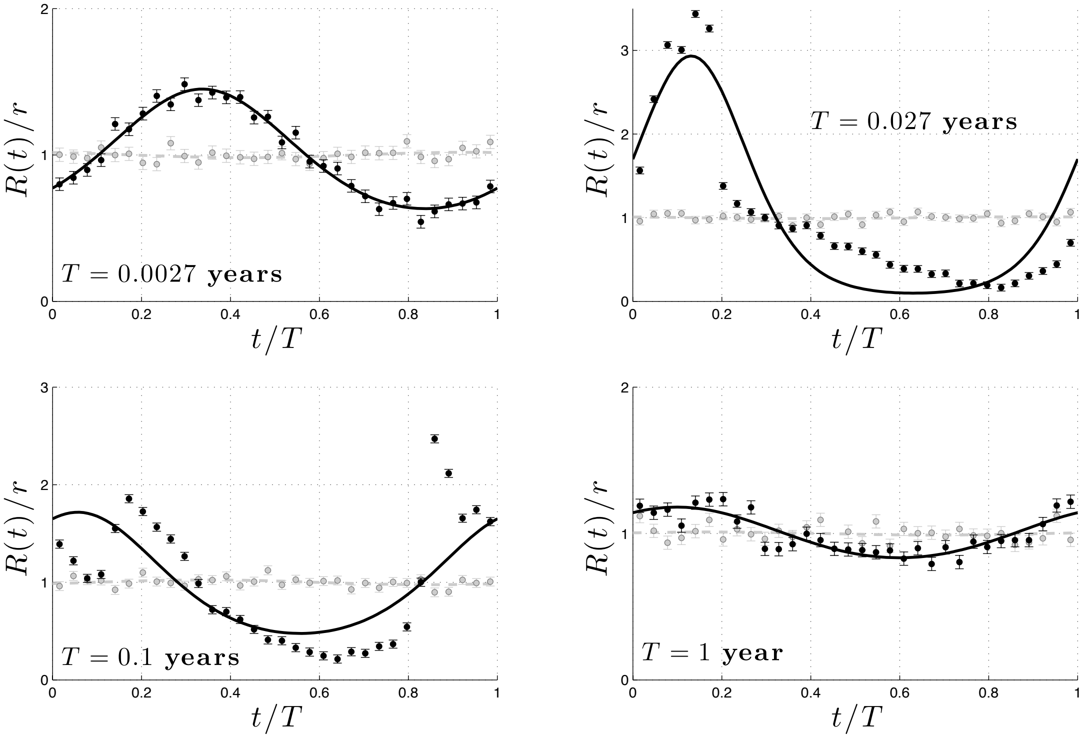


Figure 4.6: Stacked seismicity rate (black dots with error bars) and its fit with equation (4.19) (black line). The grey circles with error bars and dashed grey line show the seismicity rate from the unperturbed catalog stacked over the same period and the corresponding fit. The seismicity rate is normalized by its average, so that the quantity plotted is $R(t)/r$. The error bars on the seismicity rate only depend on the total number of events N in the catalog and the number B of bins used to compute the stacked seismicity rate: $\sigma_B = \sqrt{(B-1)/N}$ (e.g., *Ader and Avouac*, 2013). The period of perturbation T is indicated on each plot.

linear and the covariance matrix associated to the vector of parameters $[m_1 m_2]$ is:

$$C_M = 2\frac{B-1}{BN}\mathcal{I}_2, \tag{4.21}$$

where $\mathcal{I}_2$ is the 2×2 identity matrix. Given that the expected variance of the residuals is N/B (e.g., *Ader and Avouac*, 2013), in order to account for the misfit of equation 4.20 to the seismicity rate, the covariance matrix C_M is multiplied by var(residuals)$\times B/N$ if the variance of the residuals var(residuals) is greater than N/B. The uncertainties on β and Φ are finally computed from the obtained covariance matrix, which is fast because analytical. This method most likely underestimates the errors on β when the linear condition $\beta \ll 1$ is not verified, but in practice yields good orders of magnitudes for the uncertainties. In Figure 4.6, for instance, the case $T = 0.0027$ years is the most pathological and yields $\beta \approx 1.7 \pm 0.12$. The uncertainty on β is thus probably underestimated, but still provides a correct order of magnitude of the actual uncertainty, and a better estimate of this uncertainty is never needed in this study.

4.5.2 Influence of the period of the shear-stress perturbation

The procedure of section 4.5.1 is repeated for different perturbing periods (Figure 4.7). We plot both the amplitude of the seismicity rate variations (parameter β in equation (4.19)) and their phase shift (parameter Φ in equation (4.19)) with respect to the stress perturbation for the different values of the perturbation's period. The amplitude plot has been normalized by the maximum possible amplitude according to the SRM, i.e., $\beta_{\mathrm{SRM}} = \Delta\tau/a\sigma$. The predictions of the SRM are indicated for comparison as a dashed grey line, where we have taken $2\pi t_a = 0.1$ years in order to approximately fit the phase change from the simulations in Figure 4.7. This value of t_a is quite arbitrarily selected at this stage, because the secular stress loading rate $\dot{\tau_a}$ and thus t_a are not as well defined for a finite fault as for the SRM, since in the case of a finite fault, $\dot{\tau_a}$ varies both in space and time. The harmonic response of the finite fault plotted in Figure 4.7 displays features in agreement with the predictions of the SRM, but also major qualitative and quantitative differences.

There are two important differences between the response of the finite fault in our simulations and the SRM. First, the dependence of the finite-fault seismicity response on the perturbation period is non-monotonic, with a pronounced peak. We will denote the period at which the seismicity peaks by T_a. Note that the existence of such a peak may explain the differences between the response of the seismicity in the Himalaya to seasonal

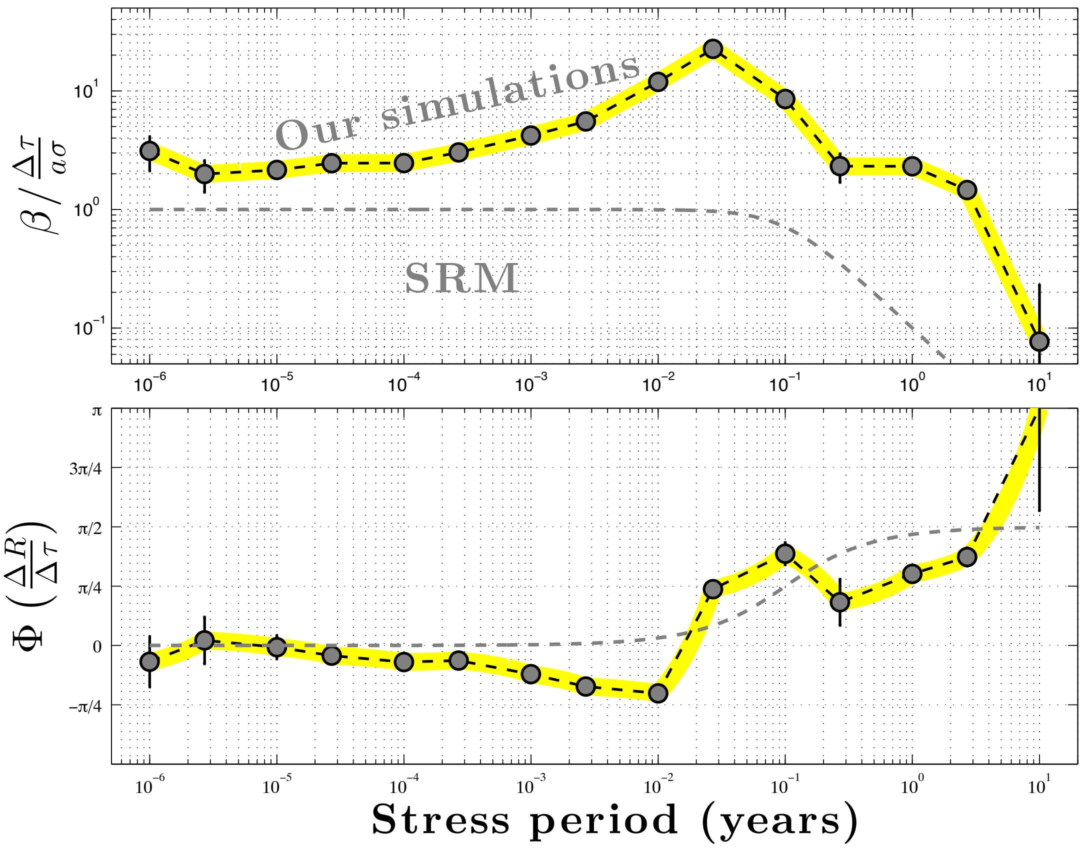


Figure 4.7: Response of a finite rate-and-state fault to harmonic shear-stress perturbations. Each point on the plot corresponds to a simulation which generated an earthquake catalog. The times of events from the catalog are stacked over one period and the resulting stacked seismicity rate is fitted with equation (4.19). The value obtained for β is reported on the upper plot, while the phase shift Φ between the seismicity rate and the stress is represented on the lower plot. Dashed grey curves show the predictions of the SRM, where we have taken $2\pi t_a = 0.1$ years in order to fit the phase. As in the SRM, one can separate two regimes of response, depending on whether the perturbing period is shorter or greater than a critical period T_a. However, the amplitude of the response is always greater than predictions from the SRM, sometimes by more than an order of magnitude, and this amplitude of the response increases with the period T for periods $T < T_a$.

perturbations and tides, provided that the period T_a is close to one year and hence the fault is more sensitive to the seasonal perturbations than perturbations of smaller periods. In Figure 4.7, $T_a \approx 0.03$ years, not yet appropriate for explaining the response in the Himalaya, and that is why we investigate the dependence of T_a on the model parameters in the following sections. Second, the amplitude of the seismicity response for the finite fault is always much larger than the predictions of the SRM. As discussed in section 4.7.3, this finding would modify how $a\sigma$ is estimated based on observations.

At the same time, as predicted by the SRM, one can distinguish between two regimes of response of the seismicity, depending whether the period T of the perturbation is larger or smaller than the critical period T_a, at which the amplitude of the seismicity response peaks. At periods $T < T_a$, variations of the seismicity rate tend to be in phase with the stress perturbation ($\Phi \sim 0$), although a gradual time lag from 0 to about $-\pi/4$ appears as the period increases, which is not expected in the SRM. On the other hand, for periods $T \geq T_a$, the seismicity rate appears to correlate with the rate of shear stress perturbation, since the phase is $\Phi \sim \pi/2$ and the amplitude seems to decrease following a $1/T$ trend. At these periods, the response of the seismicity rate seems thus analogous to the Coulomb Failure type of response, as predicted by the the SRM at periods such that $T \gg t_a$.

In order to better understand the high sensitivity of the finite fault to stress variations, we examine in more detail the seismicity produced by the perturbed fault. Figure 4.8 shows the distribution of event magnitudes produced by the harmonically perturbed fault, at periods $T = 0.0027$ years, $T = 0.027$ years and $T = 1$ year. This Figure should be compared to Figure 4.4, which shows the distribution of event magnitudes on the unperturbed fault. At period $T = 1$ year, when the fault responds in a Coulomb-like fashion (Figure 4.7), the distribution of magnitudes on the fault is quite similar. However, for the perturbation periods shorter than T_a ($T = 0.0027$ and 0.027 years), the distribution of magnitudes is considerably modified. The fault only produces either events of magnitude $M_{\text{lin}} \sim 1.05$ rupturing the entire seismogenic patch or events with $-0.4 \leq M_{\text{lin}} \leq 0.4$, i.e., smaller on average than the events usually produced by the unperturbed fault, which have magnitudes $-0.2 \leq M_{\text{lin}} \leq 0.6$. Events of magnitude $M_{\text{lin}} \sim 0.94$ that appear to form a stable pattern on the unperturbed fault have completely disappeared from the seismicity on the fault perturbed at periods $T < T_a$.

Figure 4.9 shows the Schuster spectra of the seismicity on the faults perturbed with the same three periods. Except for a prominent peak at 1 year caused by the response of the seismicity to the stress perturbation, the spectrum of the fault perturbed at 1 year

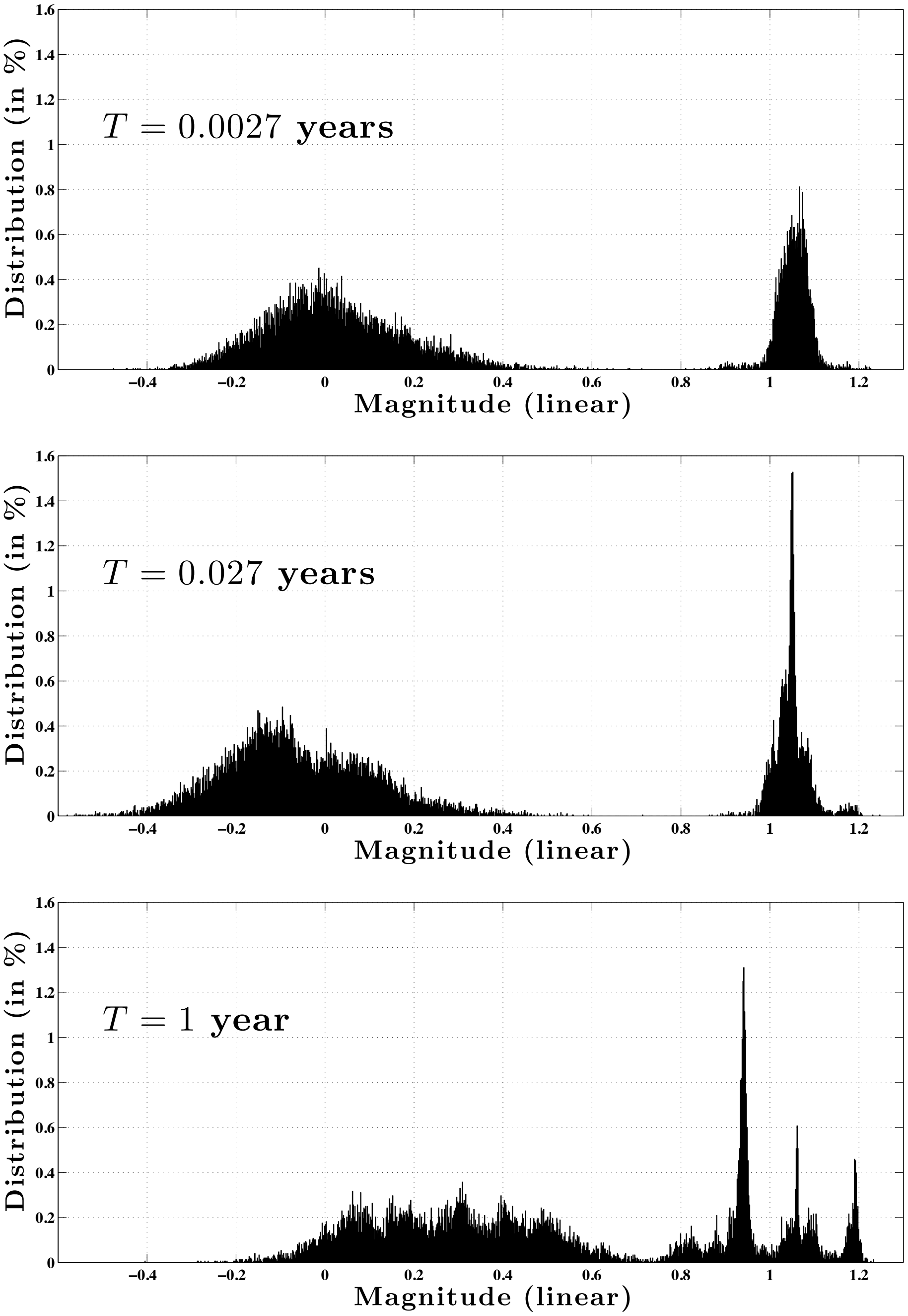


Figure 4.8: Magnitude distribution of events happening on a fault loaded at $V_{pl} = 1$ cm/yr (i.e., grey circles in Figure 4.7), for harmonic shear stress perturbations at periods $T = 0.0027$ years, $T = 0.027$ years (thus $T < T_a$) and $T = 1$ year ($T > T_a$). The other simulation parameters are the same as for Figure 4.7. While the distribution of magnitudes produced by the fault perturbed at 1 year is essentially the same as the one of the unperturbed fault (Figure 4.4), the distribution is very different for the two faults perturbed at periods $T < T_a$. The perturbed fault produces smaller events than the unperturbed fault, and events that rupture the entire seismogenic patch all have $M_{\text{lin}} \sim 1.05$. In particular, $M_{\text{lin}} \approx 0.94$ and $M_{\text{lin}} \sim 1.2$ events have almost completely disappeared from the seismic population produced by the fault.

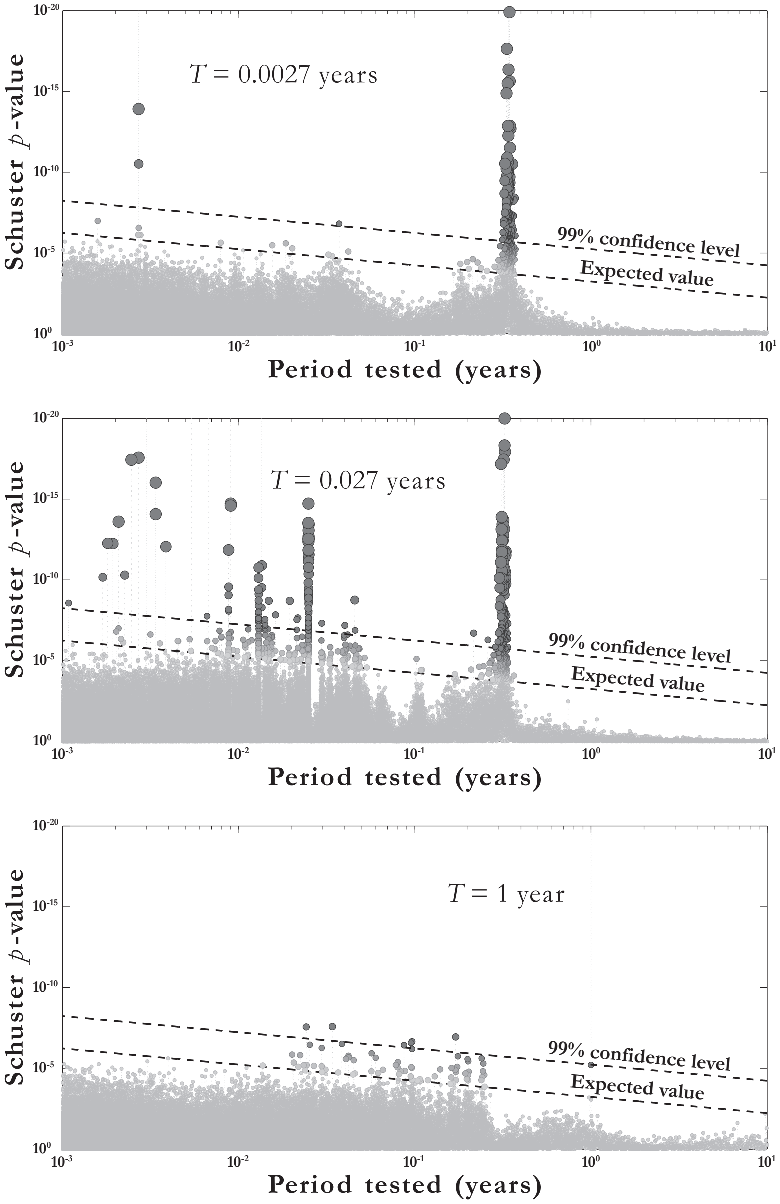


Figure 4.9: Schuster spectra for timing of events happening on a fault loaded at $V_{pl} =$ 1cm/yr (i.e., grey circles in Figure 4.7), for harmonic shear stress perturbations at periods $T = 0.0027$ yrs, $T = 0.027$ yrs and $T = 1$ yr. Other simulation parameters are the same as for Figure 4.7. We saturated the y-axis at 10^{-20} in order for the plots to be readable. For the perturbations at $T = 0.0027$ years (upper plot), the Schuster p-value at the perturbing period goes up above 10^{-150}. For the perturbations at $T = 0.027$ years (middle plot), the Schuster p-value at the perturbing period goes up to about 10^{-2150}, and all the first harmonics go up to values above 10^{-100}. For the perturbation at $T = 1$ year (lower plot), the Schuster p-value at the annual period goes up to about 10^{-50}. Except for the prominent peak at $T = 1$ year, the spectrum of the fault perturbed at 1 year (lower plot) is very similar to spectrum of the unperturbed fault in Figure 4.5. The spectra of faults perturbed at $T < T_a$ (upper and middle plot), exhibit a prominent peak at $T = 0.33$ years, which is approximately the return period of $M_{\text{lin}} = 1.05$ events, indicating that these now happen in a quasi-periodic way on the fault.

(lower plot in Figure 4.9) is essentially the same as the spectrum of the unperturbed fault in Figure 4.5. The periodicity at $T \approx 0.2$ years, corresponding to the return period of $M_{\text{lin}} \sim 0.94$ events, is still in the spectrum, although not as prominent as in the unperturbed case. The timing of these events is most likely slightly affected by the perturbation and they do not happen as periodically as in the unperturbed case. The Schuster spectra for faults perturbed at periods $T = 0.0027$ and 0.027 years are much different from the spectrum of the unperturbed case (Figure 4.5). As expected, they have prominent periodicities at the perturbing periods and at some harmonics of the perturbing period (the y-axis has been saturated, the corresponding Schuster p-values are much smaller than what can be read on the spectra), but they also exhibit very low Schuster p-values at periods around 0.33 years, which, as has been indicated earlier, approximately corresponds to the return period of $M_{\text{lin}} = 1.05$ events on the fault. As highlighted earlier, these $M_{\text{lin}} \sim 1.05$ events now constitute all of the events produced on the fault that rupture the entire seismogenic patch. Although they used to happen in a fairly chaotic way on the unperturbed fault, the strong periodicity at their return period indicates that they now occur fairly regularly. The characteristics of the seismicity on the fault are thus drastically changed when the fault undergoes harmonic stress perturbations at periods $T < T_a$, consisting exclusively of large $M_{\text{lin}} = 1.05$ events, almost systematically followed by two small aftershocks, with the rare occurrence of small events during the interseismic period (this is obvious when looking at the slip on the fault in this case, supplementary Figure S1). The perturbation seems to have stabilized the seismic cycles dominated by $M_{\text{lin}} = 1.05$ events, while the seismic regime stable in the case of an unperturbed fault (repeating $M_{\text{lin}} \sim 0.95$ events) has literally disappeared from the seismicity. It also seems that the fault is now unable to produce larger $M_{\text{lin}} \sim 1.2$ events.

This highlights a first major conceptual difference between the finite-fault simulations and the SRM, which assumes that the stress perturbation simply modifies the timing of events on the fault. Our simulations suggest that, beyond a mere change in the timing of events, the entire set of characteristics of the earthquake population produced by the fault is modified. This might explain why the response of the seismicity is much larger in the case of the finite fault than with the SRM.

The modification of the seismicity pattern by the perturbation is even more obvious when studying a fault with a smaller seismogenic patch. Figure 4.10 shows the slip on a fault of the same size (3 km), but where the seismogenic patch is 200 m long. The unperturbed fault (Figure 4.10a) only produces one type of an earthquake, happening like a clock with

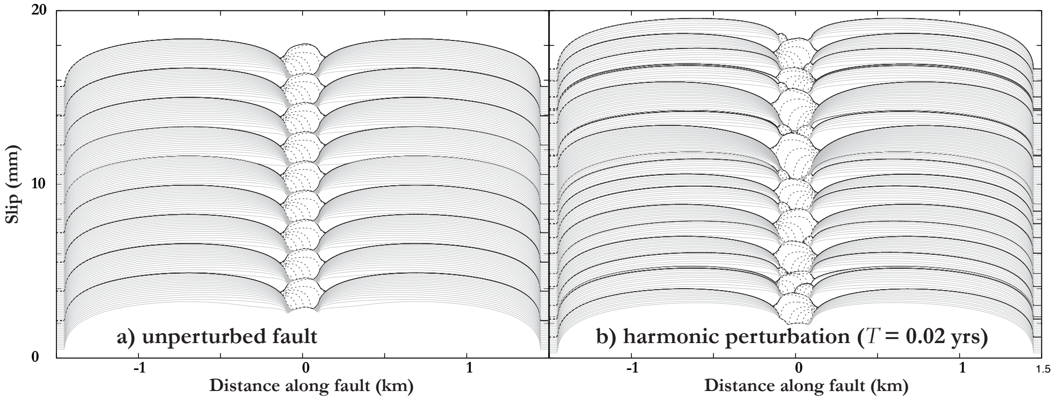


Figure 4.10: Unperturbed and perturbed slip evolution of a fault. a) Unperturbed evolution of the fault, which has the same features as the previous fault, except that the seismogenic patch is only 200 m long. In this case, the fault only produces one type of earthquake, at equally-spaced time intervals (return period of about 0.17 years). b) Slip on the fault when it is perturbed with a harmonically varying shear stress with period $T = 0.02$ years and amplitude $\Delta\tau = 3$ kPa. The seismicity produced by the perturbed fault is much different, illustrating that the perturbation may have a much broader impact on the seismicity produced by the fault than just inducing a simple modification of the timing of events.

the exact same magnitude and at constant time intervals. The seismicity produced is much less diverse than the one produced by a fault with a 500 m long patch. However, when a small harmonic perturbation is applied, the seismicity produced by the 200 m-long patch becomes diversified (Figure 4.10b). The patch starts to produce events that rupture the entire seismogenic patch and events that only happen at the edges, so that the seismicity distribution covers a larger range of magnitudes. Even more clearly here than with the 500 m patch, the introduction of an external harmonic perturbation does much more than simply affect the timing of events: it modifies the type of earthquakes that the fault can produce.

4.5.3 Influence of the background loading rate

In the SRM, the response of the seismicity to a harmonic stress perturbation depends on the relative values of the perturbation's period and the characteristic period T_a which is inversely proportional to the secular loading rate $\dot{\tau_a}$ (equation (4.8)).

In order to see how the secular loading rate affects the frequency response of the finite fault, we look at the frequency response under different loading velocities. In the case of the finite fault, the relation $\dot{\tau_a} = kV_{pl}$ is not well defined any more since $\dot{\tau_a}$ varies in space

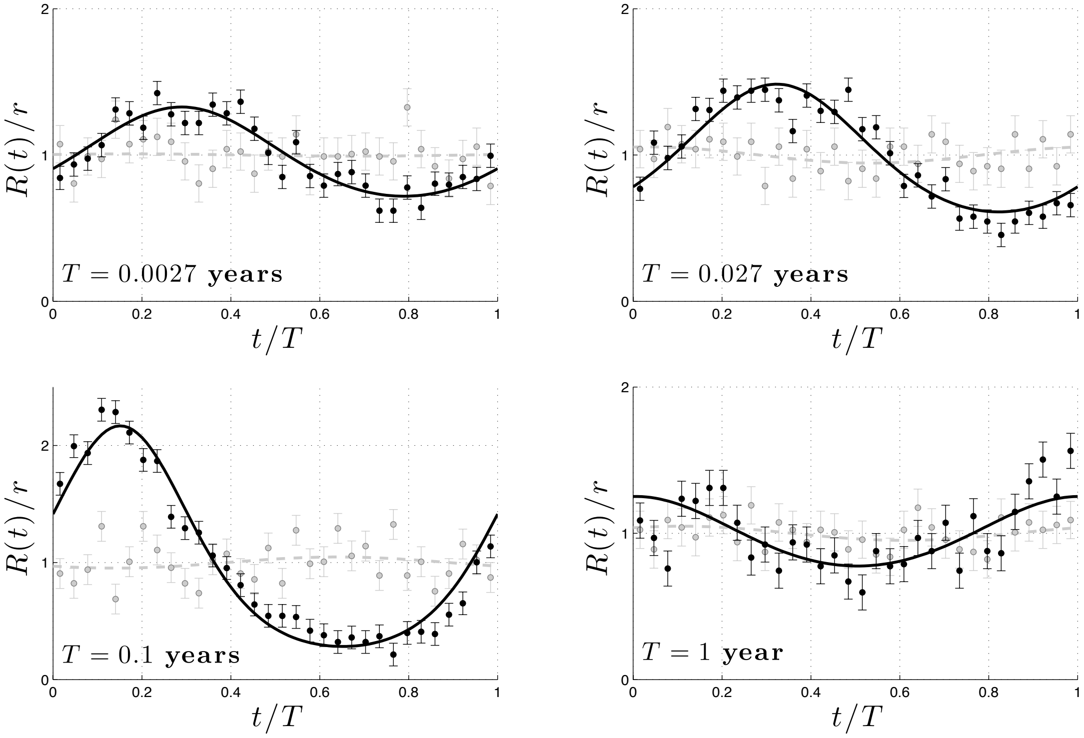


Figure 4.11: Stacked seismicity rate (black dots with error bars) and its fit with equation (4.19) (black line) for $M_{\text{lin}} > 1$ events only. This plot is similar to Figure 4.6, except that we only select $M_{\text{lin}} \geq 1$ events. Equation (4.19) fits the seismicity rate quite well, indicating that the SRM describes qualitatively well the seismicity rate on the perturbed fault for events rupturing the entire seismogenic patch.

and time. For example, when an event ruptures the entire patch, the following events often happen concurrently with the afterslip of this initial event, and therefore under a higher local $\dot{\tau}_a$. As a result, in order to single out the effect of the secular loading, we look at the response of $M_{\text{lin}} > 1$ events, for which afterslip does not have any effect anymore, and which thus happen only under the loading due to the constant loading velocity applied at the edges of the fault.

The stacked seismicity rate for these large events is represented in Figure 4.11 for the same periods as in Figure 4.6. In this case, some of the complexity disappears and equation (4.19) gives a good fit to the seismicity rate at all periods, indicating that the SRM provides a good qualitative representation of the seismicity.

Looking at the quantitative response, Figure 4.12 shows the response of $M_{\text{lin}} > 1$ events to harmonic perturbations of shear stress, for different values of the background loading velocity: the pink squares with the red line show the response for $V_{pl} = 10$ cm/yr, the grey circles with yellow line show the response for $V_{pl} = 1$ cm/yr and the blue triangles

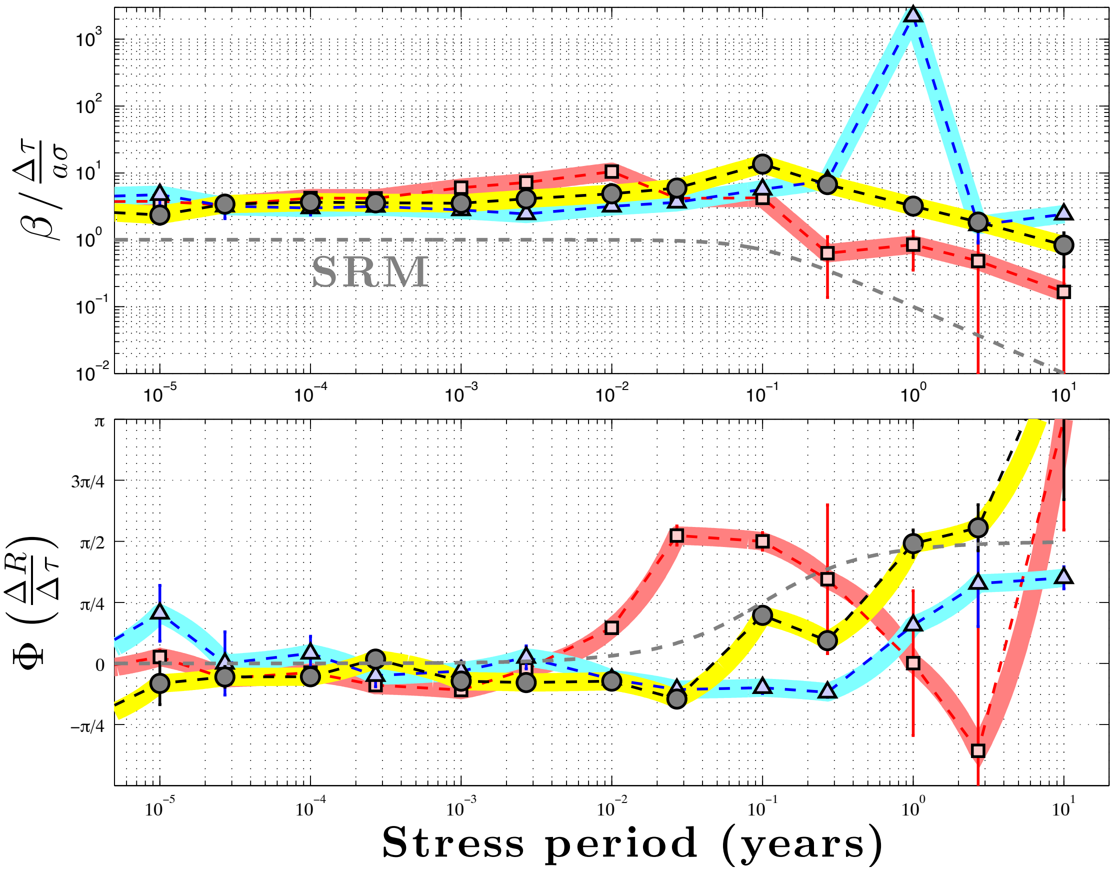


Figure 4.12: Response of a rate-and-state fault to harmonic shear-stress perturbations, for three different values of the plate loading velocity (blue triangles with the blue line: $V_{pl} = 0.1$ cm/yr, black circles with yellow line: $V_{pl} = 1$ cm/yr, pink squares with the red line: $V_{pl} = 10$ cm/yr) for $M_{\text{lin}} \geq 1$ events. Other parameters of the simulation and the method to generate the plot are the same as in Figure 4.7. The critical period T_a at which the amplitude of the response is the largest appears to be inversely proportional to V_{pl}, in qualitative accordance with predictions of the SRM. Interestingly, the value of V_{pl} has an impact on the amplitude of the frequency response, especially at the critical period T_a.

with the blue line show the response for $V_{pl} = 0.1$ cm/yr. For all these values of V_{pl}, the amplitude of the harmonic response at these magnitudes always remains much greater than the predictions of the SRM (dashed lines). In other words, if equation (4.19) provides a good qualitative representation of the variations of the seismicity rate for large events, the amplitude β of the response has to be much larger than the predictions of the SRM in order to fit the seismicity rate variations, sometimes by several orders of magnitude.

In terms of the impact of the secular loading on the response of the fault to harmonic stress perturbations, V_{pl} influences both the amplitude of the response and the critical period T_a for which this amplitude is the largest (Figure 4.12). More specifically, the critical period T_a appears to be inversely proportional to the plate velocity V_{pl}, analogously to the inverse proportionality of t_a and $\dot{\tau_a}$ predicted by the SRM. The amplitude of the response, however, is not a simple translation along the periods axis as predicted by the SRM. The amplitude at the critical period T_a seems to increase as V_{pl} gets smaller. The difference is not obvious between loading velocities $V_{pl} = 1$ and 10 cm/yr, but the peak amplitude at $V_{pl} = 0.1$ cm/yr is about 2 orders of magnitude larger than at $V_{pl} = 1$ cm/yr. Such a peak looks like a resonance at the characteristic return period of events T_{RET}, but would have no obvious reason to happen at $T = 1$ year for $V_{pl} = 0.1$ cm/yr, and not at $T = 0.1$ years for $V_{pl} = 1$ cm/yr, since $T_{\mathrm{RET}} \propto 1/V_{pl}$ (equation (4.15)).

The inverse dependence of T_a on the loading rate appears in other features of the response of seismicity. For instance, comparing the frequency response of all and $M_{\mathrm{lin}} > 1$ events for $V_{pl} = 1$ cm/yr (grey circles with yellow line respectively in Figures 4.7 and 4.12), the critical period T_a is higher for the response of $M_{\mathrm{lin}} > 1$ events. This simply comes from the fact that smaller events (i.e., events that only rupture an edge of the seismogenic patch) often occur concurrently with the afterslip of large events, and hence they have a higher loading rate than larger events, and, assuming an inverse variation of T_a with $\dot{\tau_a}$, their apparent T_a is smaller. When looking at all events together (Figure 4.7), since there are more smaller events than larger events, the global variations of seismicity rate on the fault are dominated by those of the smaller events, and therefore the apparent T_a is the one of the smaller events.

Events that rupture the entire seismogenic patch and have a magnitude $M_{\mathrm{lin}} > 0.9$ are usually followed by two aftershocks. Comparing the timing of these first two aftershocks with the stress perturbation ongoing on the fault brings to light another manifestation of the inverse dependence of T_a on the loading rate. Figure 4.13 shows the distribution of the timings of $M_{\mathrm{lin}} > 0.9$ events stacked over one perturbing period and the relative timing

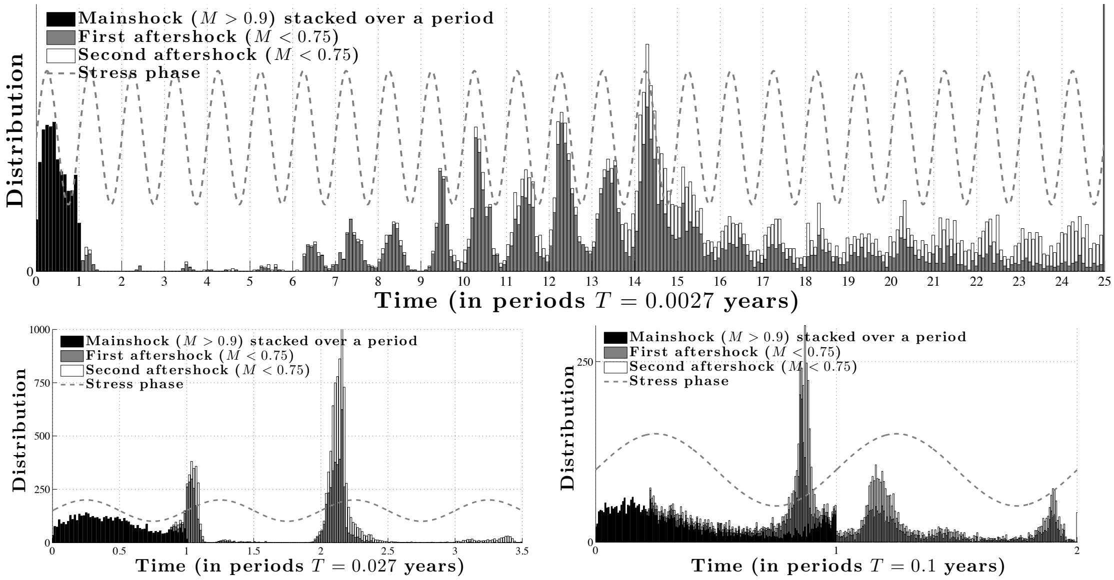


Figure 4.13: Timing of mainshocks ($M_{\text{lin}} > 0.9$) stacked over the perturbing period ($T = 0.027$ yrs) for a fault loaded at $V_{pl} = 1$ cm/yr, and relative timing of the first two small aftershocks ($M_{\text{lin}} < 0.75$). The time distribution of the mainshocks is indicated with black bars, while the one of the aftershocks is indicated with grey bars. The dashed grey sine-wave schematically shows the phase of the stress applied on the fault. The timing of aftershocks seem largely affected by the stress perturbation and their phase with respect to the stress perturbation varies through time, owing to the influence of the afterslip caused by the mainshock.

of the first two aftershocks. Note that $M_{\text{lin}} > 0.9$ events are about the same set as the $M_{\text{lin}} > 1$ events, as the distribution of magnitudes from Figure 4.8 indicates that very few events have magnitudes between 0.9 and 1. After a $M_{\text{lin}} > 0.9$ event happens, afterslip at the edge of the seismogenic patch induces the stressing rate $\dot{\tau}_a$ to suddenly increase, and then slowly decay back to its interseismic value (e.g., *Perfettini and Avouac*, 2004). The inverse dependence of T_a with the loading rate would thus cause T_a to suddenly drop after the mainshock (i.e., the $M_{\text{lin}} > 0.9$ event) and then increase back to its interseismic value.

With this expected variation of T_a in mind, the variation of phase of aftershocks in the case of a perturbing period $T = 2.7 \times 10^{-3}$ years (Figure 4.13 upper plot) can be understood. Figure 4.12 (grey circles with yellow line) indicates that the phase Φ between the seismicity rate of large events and the shear stress perturbation is slightly negative, meaning that the peak of seismicity rate happens slightly after the peak of stress when no afterslip occurs. However, Figure 4.13 (upper plot, $T = 0.0027$ years) shows that if the mainshocks have an expected slightly negative phase, the aftershocks have different phases depending on their timing relative to the mainshock. Aftershocks happening within one period of the mainshock have a positive phase $\Phi > 0$ (they happen before the maximum of the shear stress, almost at the maximum of shear stress rate), while the phase of aftershocks happening around 9 or 10 periods later drops to $\Phi \approx -\pi/4$ and then gradually increases to $\Phi \sim 0$ for aftershocks happening 14 or more periods after the mainshock. Noting from Figure 4.12 that $\Phi > 0$ corresponds to $T_a < T$, $\Phi \approx -\pi/4$ to $T_a \gtrsim T$ and $\Phi \sim 0$ to $T_a \gg T$, this variation of phase can indeed be explained by the value of T_a dropping right after the mainshock and then slowly increasing, while the period T of the stress perturbation remains the same.

The case of the perturbing period $T = 0.027$ years (Figure 4.13 middle plot) displays a similar pattern. The mainshocks happen with a slightly negative phase ($T < T_a$ according to Figure 4.7), and the aftershocks happen within either 1 or 2 periods of the mainshock. Aftershocks happening within one period of the mainshock have an almost $\pi/2$ phase (Coulomb regime, $T_a \ll T$), while the phase of the ones happening during the next period starts decreasing ($T_a \lesssim T$).

Figure 4.13 also shows the case of a perturbing period $T = 0.1$ years (lower plot), which illustrates the complexity of the seismicity on the fault. The mainshocks are getting close to the Coulomb regime ($\Phi > 0$ in Figure 4.12 for $T = 0.1$ years), and so the aftershocks would tend to remain in the same regime when decreasing T_a as $\dot{\tau}_a$ increases due to afterslip. However, in this case, the period of the perturbation becomes larger than the typical mainshock-aftershock time, which is usually less than 0.05 years, as can be seen in the cases of perturbations periods $T = 0.0027$ and $T = 0.027$ years. As a result, when mainshocks happen at the beginning of the period, which is the case for the majority of them, induced aftershocks cannot "wait" an entire period to happen, and they nucleate before the end of the period as soon as the stress on the fault starts increasing again. The global shear stress reduction due to the perturbation thus still appears to prevent them from nucleating, but as soon as the perturbing stress starts increasing again, these events nucleate, resulting in a peak in the seismicity rate at the end of the period. These considerations explain the complexity that is observed in the seismicity rate in Figure 4.6 for periods $T = 0.027$ and 0.1 years: the different peaks in the seismicity rate within one period correspond to different populations of events that happen under different loading conditions, and thus with different phases. Stacking them altogether thus results in a complex seismicity rate that equation (4.19) thus cannot fit properly. This explains why isolating large events as is done in Figure 4.11, which all respond with the same phase, leads to a much more successful fit of equation (4.19) to the seismicity rate.

4.5.4 Influence of the characteristic rate-and-state slip D_c

Studying the influence of D_c is challenging, because varying the value of D_c on the fault has several implications. When D_c increases, equation (4.13) indicates that the nucleation size of events increases proportionally. When the nucleation size is not small enough compared to the length of the seismogenic patch (typically, when the nucleation size becomes larger than a tenth of the length of the seismogenic patch), only events of similar magnitude rupturing the whole patch occur. Such a regular seismicity has a different behavior, with resonances appearing between the return period of events produced and the period of the perturbation, and comparing the response of such a fault to the response of a fault able to produce a wide range of magnitudes becomes difficult. On the other hand, when D_c decreases, the size of the cohesion zone defined in equation (4.10) decreases proportionally, and the cell size for the simulation has to be reduced accordingly, which requires longer simulation times. Therefore, the parameter D_c can only be varied within a limited range of values.

Nevertheless, equations (4.10) and (4.13) indicate that both the size of the cohesion zone and the nucleation size are proportional to D_c/σ. In theory, varying D_c and σ so that the ratio D_c/σ remains constant should therefore result in both a constant nucleation size and a constant size of the cohesion zone, i.e., a constant cell size required to simulate the evolution of slip on the fault. Since such simulations alone cannot assess whether changes in the fault response are due to variations of D_c or σ, we therefore run simulations where we vary only D_c, simulations where we vary only σ, and simulations where we vary both D_c and σ together keeping the ratio D_c/σ constant. The last type of simulations lets us vary D_c over a wider range of values than the first two types of simulations.

First, we look at the response of the fault to harmonic perturbations for a reduced range of values of D_c ($D_c = 1$, 2, 5 and 10 μm) keeping the other parameters constant. When varying D_c, the characteristic period T_a at which the amplitude of the seismicity-rate variations is the largest seems to remain approximately proportional to D_c (Figure 4.14). Figure 4.14 also seems to indicate that the amplitude of the variations of the seismicity rate at perturbing period T_a increases with D_c, although this could be due to a period-sampling effect, i.e., the periods chosen for simulation.

The case of $D_c = 10$ μm (blue triangles in Figure 4.14) is presented in Figure 4.14 in order to show the complexity and the nonlinearity of the response, but actually does not provide much insight about the value of T_a. In this case, the nucleation size is too large for the unperturbed fault to produce a variety of magnitudes, and only $M_{\text{lin}} = 1.09$ event are produced at the regular return period of 0.38 years. Some complexity happens when the fault is perturbed at periods smaller than this regular return period, similarly to the fault shown in Figure 4.10. But when the fault is perturbed around the return period of events, the timing of seismic events starts to resonate with the perturbation and the amplitude of the response becomes quite large. The peak at 0.3 years in Figure 4.14 is due to this resonance and precludes us from determining T_a and the associated peak for this set of parameters if such resonance did not exist (i.e., for a longer fault segment).

Note that in the cases of $D_c = 1$, 2 and 5 μm, the period T_a is smaller than the characteristic return period of events considered on the fault. For $D_c = 5$ μm, the return period of events rupturing the whole patch (i.e., $M_{\text{lin}} \approx 1.05$ events) is about 0.31 years and $T_a \approx 0.08$ years, while for $D_c = 1$ and 2 μm, the return period of the considered events is about 0.15 years ($M_{\text{lin}} \approx 0.85$) whereas $T_a \approx 0.02$ or 0.03 years. In these cases, the increase of the response amplitude is thus not due to a resonance with the natural return period of events. The fact that a resonance at the return period of considered events will impair the determination of T_a is also illustrated in Figure S8 in the supplementary material, where we compare the response of a fault with $D_c = 1$ μm and $\sigma = 1$ MPa to stress perturbations of amplitude $\Delta\tau = 0.6$ kPa and $\Delta\tau = 3$ kPa. Figure S8 is coincidently a good illustration of the nonlinearity of the response of the finite fault with $\Delta\tau$.

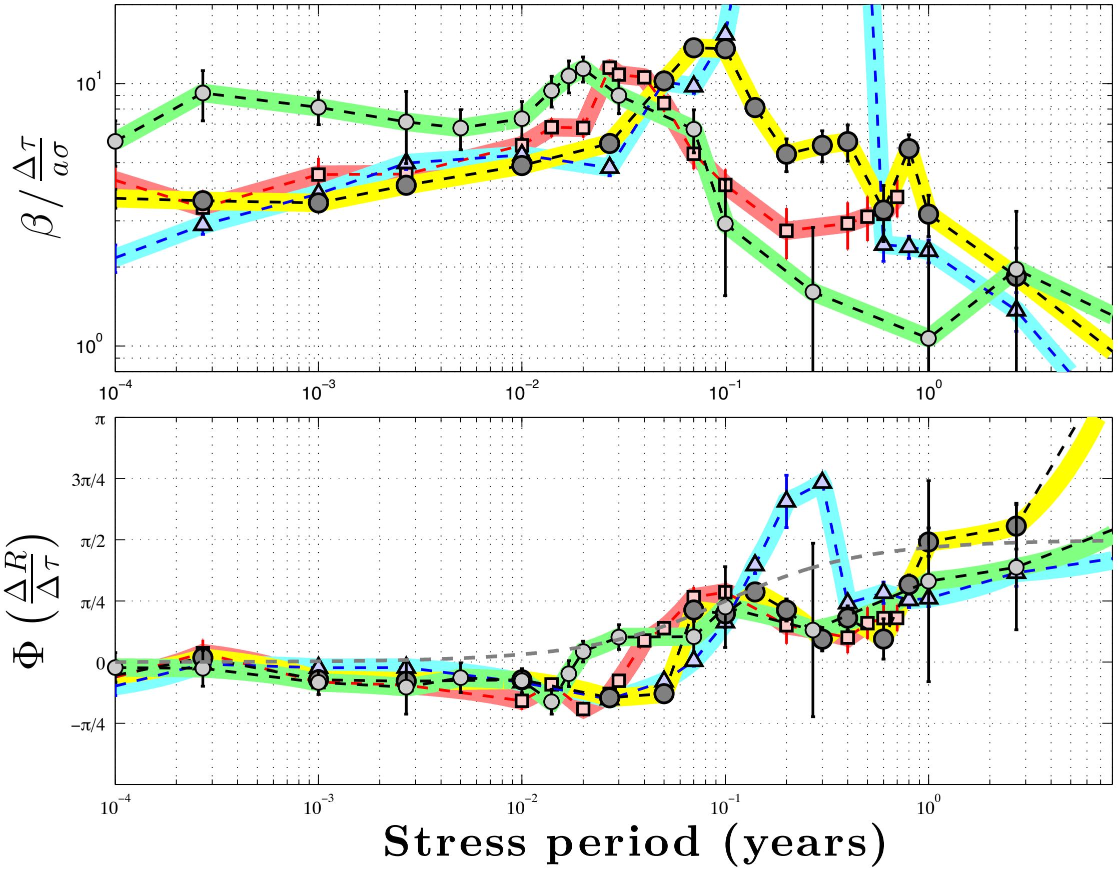


Figure 4.14: Response of a finite rate-and-state fault to harmonic shear-stress perturbations for four different values of the fault parameter D_c, for events rupturing the whole seismogenic patch. Other parameters are the same as in Figure 4.7. Blue triangles with the blue line: $D_c = 10\ \mu$m; black circles with yellow line: $D_c = 5\ \mu$m; pink squares with the red line: $D_c = 2\ \mu$m; light grey circles with green line: $D_c = 1\ \mu$m. The critical period T_a at which the amplitude of the response is maximum seems to be approximately proportional to D_c. The peak in amplitude in the case of $D_c = 10\ \mu$m is due to a resonance at the return period of typical events produced by the fault, which makes it dependent on the seismogenic segment size and hence different from T_a dependencies in other cases.

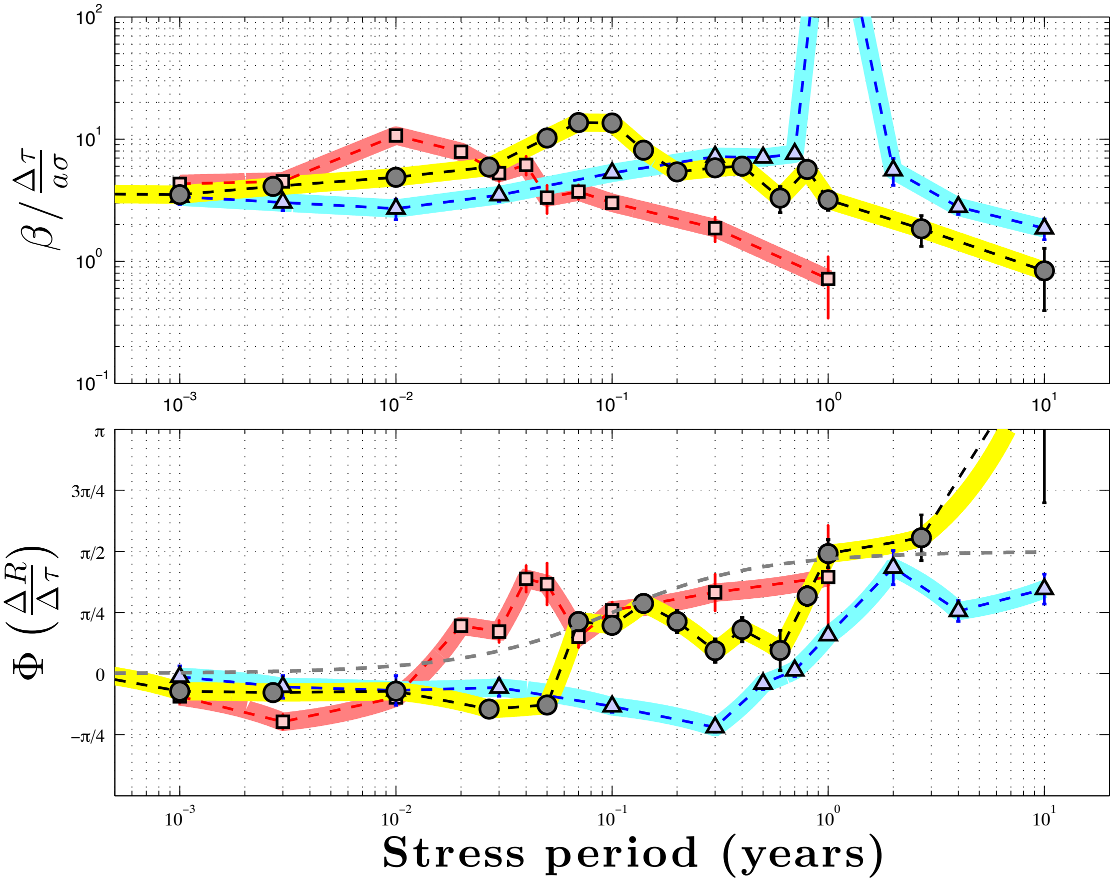


Figure 4.15: Response of a finite rate-and-state fault to harmonic shear-stress perturbations for three different values of the fault parameter D_c, as well as varying σ and $\Delta\tau$ to keep ratios D_c/σ and $\Delta\tau/a\sigma$ constant for the three simulations. Only events rupturing the whole seismogenic patch are used to compute the variations of seismicity rate. Blue triangles with the blue line: $D_c = 50$ μm; black circles with yellow line: $D_c = 5$ μm; pink squares with the red line: $D_c = 1$ μm. Other parameters of the simulation and the method to generate the plot are the same as in Figure 4.7. Similarly to Figure 4.14, the period T_a at which the amplitude of the response is maximum seems to be proportional to D_c.

Now that the influence of D_c alone has been examined, we vary D_c and σ together keeping the ratio D_c/σ and thus the nucleation size constant. In this case, we also vary the amplitude of the stress perturbation $\Delta\tau$ from one set of parameters to the other so that the ratio $\Delta\tau/a\sigma$ remains constant. Doing so over a wider range of values for D_c than previously explored ($D_c = 1$, 5 and 50 μm), the position of the characteristic period T_a seems again to be proportional to D_c, indicating that T_a varies with D_c but not with σ (Figure 4.15). Besides, as before, the amplitude of the variations of seismicity rate at T_a seems to increase as D_c increases.

It is noteworthy that the phase associated with this response (lower plot in Figure 4.15) seems to display anomalies, especially obvious around $T = 0.05$ years for $D_c = 1$ μm and around $T = 0.3$ years for $D_c = 5$ μm. In the case $D_c = 1$ μm and thus $\sigma = 1$ MPa, the larger events produced by the fault have a magnitude of $M_{\text{lin}} \sim 0.5$ which would have a return period of $T(M_{\text{lin}} \sim 0.5) \approx 0.05$ years, according to equation (4.16). Similarly, as has been mentioned earlier, the period $T = 0.3$ years corresponds to the return period of large characteristic $M_{\text{lin}} \sim 1.05$ events produced by the fault when $D_c = 5$ μm and thus $\sigma = 5$ MPa. These anomalies in phase are thus most likely due to a resonance between the perturbing period and the natural return period of the considered events. This resonance can also be detected at a much lower level on the amplitude plot, without affecting the general trend of the response. This is why it is preferable to select the values of T_a from the amplitude response curve.

Finally, in order to validate the assessment that the critical period T_a does not depend on the normal stress σ, we simulate the response of the finite fault for different values of σ, holding all other parameters constant. As equations (4.10) and (4.13) suggest, this exposes us to the same challenges as varying D_c in terms of the impact on the nucleation size and the size of the cohesion zone. Figure 4.16 displays again the response of the fault with $\sigma = 5$ MPa (dark grey circles with yellow line) together with the response when we increase the normal stress to 12.5 and 25 MPa (respectively pink squares with the red line and light grey circles with green line). In terms of the impact on the nucleation size, these two values of the normal stress are respectively equivalent to $D_c = 2$ and 1 μm (respectively pink squares with the red line and light grey circles with green line in Figure 4.14). Whereas decreasing D_c to 1 or 2 μm proportionally decreases the critical period T_a (Figure 4.14), increasing σ to either 12.5 or even 25 MPa does not have any systematic effect on T_a.

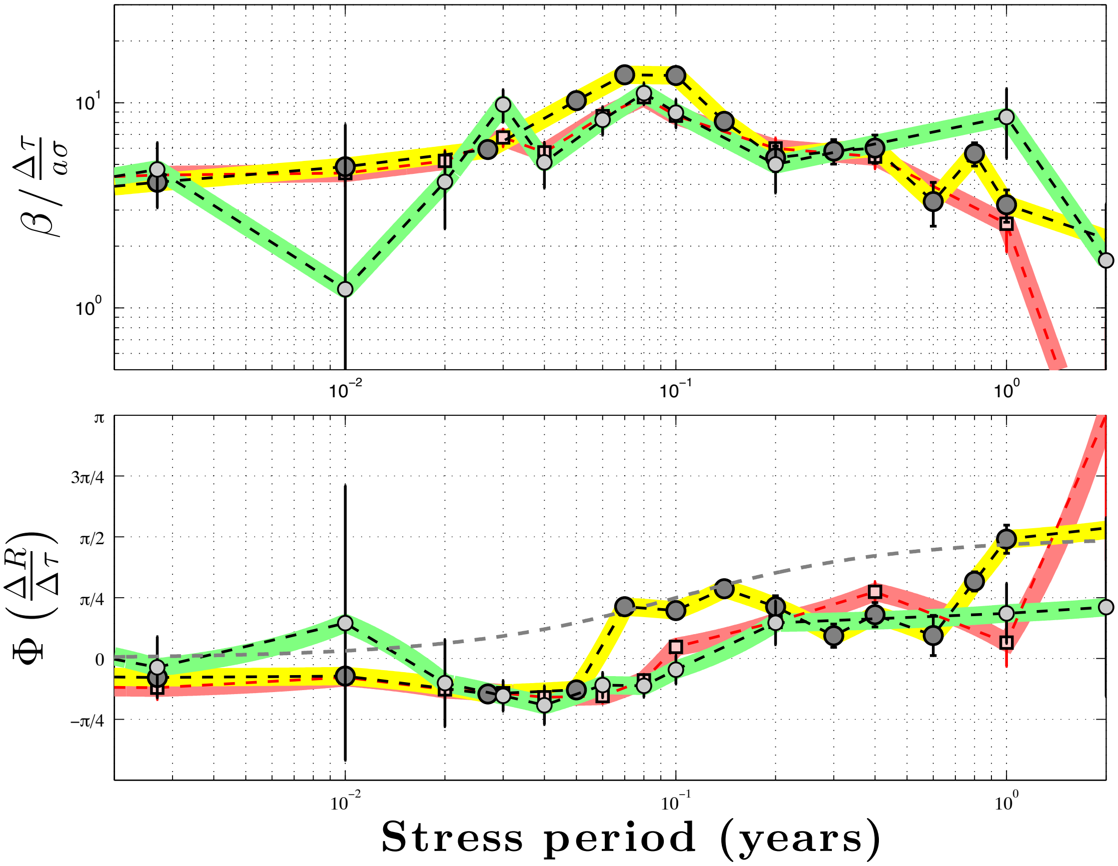


Figure 4.16: Response of a finite rate-and-state fault to harmonic shear-stress perturbations for three different values of the normal stress σ, for events rupturing the whole seismogenic patch. Large dark grey circles with yellow line: $\sigma = 5$ MPa, pink squares with the red line: $\sigma = 12.5$ MPa, small light grey circles with green line: $\sigma = 25$ MPa. The last two values of σ are respectively equivalent to $D_c = 2$ μm (pink squares with the red line in Figure 4.14) and $D_c = 1$ μm (small light grey circles with green line in Figure 4.14) in terms of changing the nucleation size and the size of the cohesion zone. As for Figure 4.14, the grid size has been changed accordingly. Unlike D_c, the normal stress does not seem to change the position of the critical period T_a.

4.6 Step response of a rate-and-state seismogenic fault

If the response of the seismicity to a shear stress history $\tau(t)$ applied on the fault was linear with the stress rate, the corresponding seismicity rate on the fault $R(t)$ could be written as a simple convolution:

$$R(t) = \mathcal{G}(t) * \dot{\tau}(t), \tag{4.22}$$

where the Green's function $\mathcal{G}(t)$ would be the response of the seismicity to a Dirac function for the stress rate, i.e., to a step function in stress. In other words, $\mathcal{G}(t)$ would simply be the seismicity rate of an aftershock sequence following a stress step of unit amplitude.

In this section, we therefore examine the response of the finite fault to a step-like perturbation in stress, and see to which extent it compares with the harmonic response of the fault described in the previous section. We do not study the step response in as much detail as the harmonic response. The goal of this section is to highlight some of the inherent properties of the finite fault revealed in the previous section and to also illustrate the nonlinearity of the response of the fault to stress perturbations.

In order to determine the response of a finite fault to a step-like perturbation of stress, we consider the fault presented in section 4.4 and alternatively impose steps of shear stress of amplitude $\Delta\tau = 40$ kPa and -40 kPa so that $|\Delta\tau/a\sigma| = 1$. With this mode of perturbation, the mean of the stress perturbation is zero in the long term, so that no shear stress accumulates on the fault. As in the previous section, the fault is loaded at three different plate velocities: $V_{pl} = 0.1$, 1 and 10 cm/yr. When the fault is loaded at $V_{pl} = 0.1$ cm/yr, the stress steps are applied every 50 years, whereas they are applied every 5 years when $V_{pl} = 1$ or 10 cm/yr. These time intervals are chosen because they are much longer than the characteristic time of response of the seismicity to a stress perturbation (e.g., Figure 4.12), so that the transitory response of the seismicity to the step is over when the next step is imposed on the fault. Besides, Figure 4.5 shows that 5 years does not correspond to any natural periodicity of the fault loaded at $V_{pl} = 1$ cm/yr, so that the steps happen at random times within the seismic cycle of the fault. This remains true for the two other loading velocities studied ($V_{pl} = 0.1$ and 10 cm/yr). This can also be verified a posteriori by checking that the distribution of the time intervals between the last event and the stress step has an exponential distribution (see supplementary Figure S5 for the case $V_{pl} = 1$ cm/yr).

Given that the stress steps are imposed at random times during the interseismic period, in order to study the response of a population of faults that would be at different stages of their interseismic cycle, we stack the timing of events happening around the positive stress steps, taking the time of the stress step as an origin of times. For instance, in the case of the fault loaded at $V_{pl} = 1$ cm/yr, we select events happening between 0.2 years before and 1 year after the stress step is imposed. This time window is multiplied or divided by 10 when V_{pl} is respectively divided or multiplied by 10, and it has been chosen in regard of the results of the frequency response of the fault, where we determined that the characteristic period was of the order of $T_a = 2\pi t_a \approx 0.1$ years (Figure 4.12), which would correspond to a characteristic time of response to a stress-step perturbation of $t_a \approx 0.02$ years.

As described in section 4.3, the SRM by *Dieterich* (1994) determines the cumulative number of events in response to a step of shear stress happening at time $t = 0$ for a spring-slider model under rate-and-state rheology. Similarly as in the case of the response to a harmonic variations of stress, we fit the cumulative number of events produced by the finite fault with an expression analogous to the prediction of the SRM (equation (4.6)):

$$\frac{N(t)}{r} = t + t_a \ln\left[e^{\beta} + \left(1 - e^{\beta}\right) e^{-t/t_a}\right] \mathcal{H}(t), \tag{4.23}$$

where the parameters β and t_a are determined to fit the results of the simulation. In the SRM, equation (4.6) indicates that $\beta_{\mathrm{SRM}} = \Delta\tau/a\sigma$ and $t_a|_{\mathrm{SRM}} = a\sigma/\dot{\tau}_a$. Besides, according to equation (4.23), the total number of events in the aftershock sequence is:

$$\frac{N_a}{r} = \beta t_a, \tag{4.24}$$

which in the case of the SRM reduces to

$$\left.\frac{N_a}{r}\right|_{\mathrm{SRM}} = \frac{\Delta\tau}{\dot{\tau}_a}. \tag{4.25}$$

Figure 4.17 shows the cumulative number of events before and after the stress step is imposed (black curve), where the ranges plotted on the axes have been scaled by a factor inversely proportional to V_{pl}. It also shows the fit of equation (4.23) to the cumulative number of events using β as a free parameter (dashed dark-grey curve), or imposing the value $\beta = \beta_{\mathrm{SRM}} = \Delta\tau/a\sigma = 1$ (dashed light-grey curve). Similarly to the case of the response to a harmonic stress perturbation, the SRM qualitatively reproduces the behavior of the finite fault but underestimates the amplitude of the response. In other words, equation (4.23) fits well the cumulative seismicity rate in Figure 4.17, but the best-fit value of β is much larger than what the SRM would predict. Imposing $\beta = \beta_{\mathrm{SRM}} = \Delta\tau/a\sigma$ as prescribed by the SRM and solving only for t_a never leads to a good fit (the best fitting curve with β_{SRM}

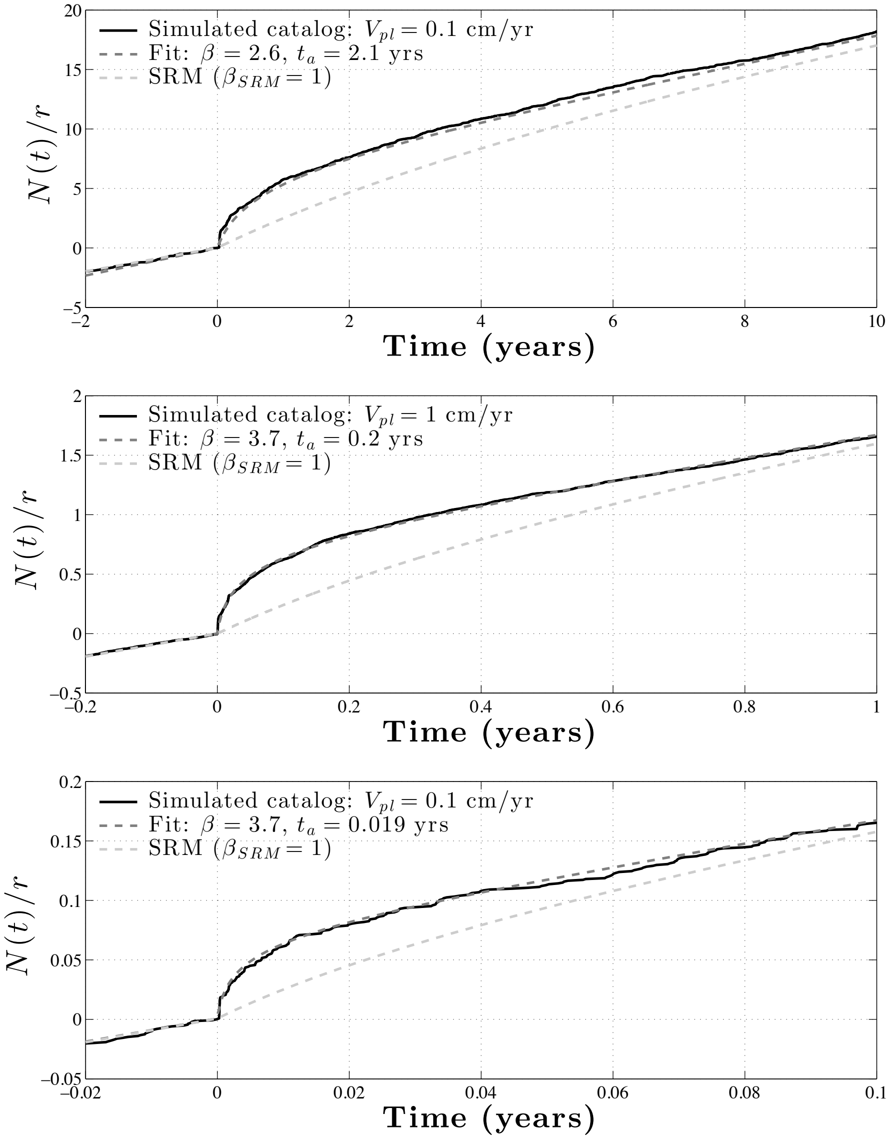


Figure 4.17: Cumulative number of events with time in response to a step of shear stress of amplitude $\Delta\tau/a\sigma = 1$ at time $t = 0$, for three values of the loading rate ($V_{pl} = 0.1$, 1 and 10 cm/yr). The other parameters of the simulations are the same as in section 4.4. Black curves show the cumulative number of events from the simulation, the dashed dark-grey curve shows the fit of equation (4.23) to the cumulative distribution of events, where parameters r, t_a and β are determined to fit the results of the simulations. The dashed light-grey curve shows the fit with equation (4.23) imposing the β-value prescribed by the SRM (i.e., $\beta_{\mathrm{SRM}} = 1$ in this case) and fitting for the best value of t_a. Note that the ranges plotted on both axes are scaled by a factor proportional to $1/V_{pl}$. As for the harmonic response, the equation predicted by the SRM qualitatively fits the response of the finite fault, but underestimates the amplitude of the response (parameter β).

is the dashed light-grey curve in Figure 4.17).

The value of t_a obtained from the fit of equation (4.23) to the cumulative number of events in the aftershock sequence generated with the finite-fault simulations ($t_a = 0.2$ years in the case $V_{pl} = 1$ cm/yr) is 10 times larger than the one computed in the case of the harmonic response ($t_a = 0.01$ years). This points out the nonlinearity of the response of a fault to a stress perturbation, and therefore the challenge underlying the quest of a general simple law linking the stress history on a fault and the resulting seismicity rate.

The characteristic time of response of the seismicity t_a is found to be inversely proportional to the loading velocity V_{pl} (Figure 4.17), as in the case of the response to a harmonic perturbation and as predicted by the SRM. This finding also concurs with the observations of *Toda et al.* (2002), who reported that the characteristic duration of aftershock sequences in the Izu islands had decreased by a factor ~ 1000 during dyke intrusions that increased the background stressing rate by a factor ~ 1000. Besides, as for the response to a harmonic perturbation, varying V_{pl} has an impact on the amplitude β of the response. However, the relation between amplitude β and background velocity V_{pl} seems different in both cases: the amplitude β of the response here appears to increase when V_{pl} increases, whereas β decreases when V_{pl} increases in the harmonic response. This points out another nonlinearity of the response of the fault to stress perturbations, indicating that looking for a linear relation such as the one described in equation (4.22) would be bound to fail.

Recall that in the case of the periodic perturbation, the magnitude distribution of events was different in the unperturbed and perturbed cases. Let us explore this phenomenon here. Figure 4.18 shows the distribution of magnitudes of the first three aftershocks happening after each stress step, together with the magnitudes of the background seismicity on the same fault and the magnitudes of events produced by an unperturbed fault, in the case of $V_{pl} = 1$ cm/yr. The background seismicity is defined here as the seismicity between 1 and 4 years after the stress step has been applied on the fault (recall that the stress steps are applied every 5 years), a time window chosen so that the effect of the stress step has died off. As expected, the background seismicity on the perturbed fault is the same as the seismicity on the unperturbed fault, but aftershocks induced by the stress step have a very different distribution of magnitudes. Some of these aftershocks have much smaller magnitudes than the unperturbed fault can produce (magnitudes between -0.6 and -0.2), while conversely, the proportion of aftershocks of magnitude $M_{\text{lin}} > 1$ is higher than the number of $M_{\text{lin}} > 1$ events on the unperturbed fault. It is noteworthy that while applying a step-like stress perturbation or a harmonic stress perturbation causes a change in the

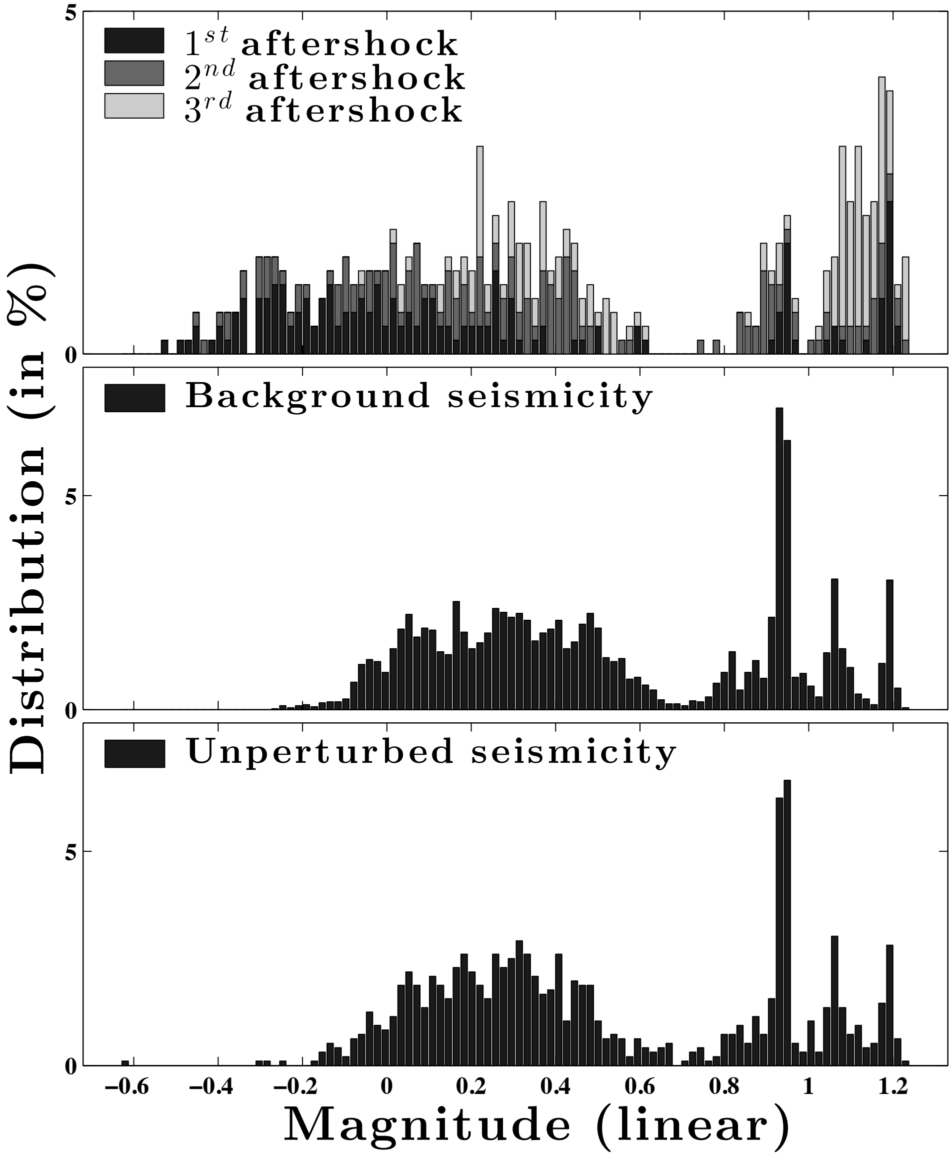


Figure 4.18: Magnitude distribution (top) of the first three aftershocks after the fault undergoes a stress step, (middle) of the background seismicity, i.e., the seismicity starting one year after the stress step has been applied, a time long enough for the the stress step not to have any effect anymore, and (bottom) of events from the unperturbed catalog. The fault has the same properties as in Figure 4.17, and is loaded at $V_{pl} = 1$ cm/yr. As for the response to a harmonic perturbation, the seismicity on the fault perturbed by a stress step is quite different from the one generated by the fault evolving unperturbed.

distribution of magnitudes produced by the fault, applying a random stress perturbation on the fault does not modify the type of events produced (Figure S6).

The results of the step response of the finite fault are thus quite similar to the ones for the harmonic response, and lead again to the major conclusion that the finite fault is much more sensitive to stress perturbations than the SRM predicts.

4.7 Discussion

4.7.1 Critical perturbation period on a finite fault

A major analogy between the finite-fault and the SRM is the existence of two regimes for the response to harmonic stress perturbations: the seismicity correlates with the stress perturbation at periods T shorter than T_a, whereas it correlates with the stress-rate perturbation at periods larger than T_a. In the SRM, the characteristic period separating the two regimes is $T_a = 2\pi t_a = 2\pi a\sigma/\dot{\tau_a}$, where the secular shear stress is $\dot{\tau_a} = kV_{pl}$, k being the spring stiffness. The fact that T_a seems to be proportional to $1/V_{pl}$ in the finite-fault model suggests that this relation should be to some extent transferable to the finite-fault model. The main issue in doing so is the fact that whereas k and $\dot{\tau_a}$ are constant in space and time and thus well defined in the SRM, they are not in the finite-fault model: depending on where and when on the seismogenic patch one measures the shear-stress rate, the value varies a lot (Figure S2).

As has been pointed out earlier, the fact that an event correlates in time with the stress perturbation is determined by the sensitivity of the nucleation zone at the onset of nucleation. In order to derive the expression of the characteristic period T_a, we therefore consider for $\dot{\tau_a}$ what would be the shear stress rate on the fault at the onset of rupture, when the nucleation zone reaches its critical size. Writing $\dot{\tau_a} = k_{ff}V_{pl}$, where k_{ff} would be the equivalent stiffness of the finite fault, $k_{ff} = G/H$, where G is the shear modulus and H the characteristic length over which interseismic deformation is accumulated.

Taking for H the whole length of the rate strengthening zone leads to $\dot{\tau_a} = 250$ kPa/yr, which is much less than any shear stress rate measured on the seismogenic patch during the nucleation (the smallest value is 1250 kPa/yr, Figure S2). Also, decreasing the size of the creeping zone by a factor of 2 does not modify the response of the seismogenic patch to a harmonic perturbation (see Figure S4 for a comparison when the size of the creeping zone is divided by 2), indicating that as long as the rate-strengthening zone is large enough to prevent any boundary effects, its actual size does not influence the behavior of the seismogenic patch.

Instead, one can assume that the characteristic length over which interseismic deformation is accumulated is of the order of a few times the critical nucleation size (*Kaneko and Lapusta*, 2008): $H = \kappa h^*$, where h^* is the critical nucleation size defined in equation (4.13) and κ a factor of the order of a few units. In this case, the critical period $T_a = 2\pi t_a = 2\pi a\sigma/\dot{\tau_a}$ is

$$T_a = 2\pi\kappa \frac{a}{F(a,b)} \frac{D_c}{V_{pl}}. \tag{4.26}$$

Taking for $F(a,b)$ the expression by *Rubin and Ampuero* (2005) leads to a background shear-stress rate $\dot{\tau_a} = 13.3/\kappa$ MPa/yr, i.e., $\kappa \sim 10$ given that in the nucleation zone $\dot{\tau_a} \approx 1.25$ MPa/yr at the crack tip at the onset of nucleation (supplementary Figure S2). With this value of κ, equation (4.26) yields $T_a = 0.2$ years, which is of the order of magnitude of what can be read in Figure 4.12. Equation (4.26) also reproduces the dependence on D_c/V_{pl} described in the section 4.5, and as expected does not a priori depend on the normal stress σ. It is possible that the prefactor κ slightly depends on σ, but this dependence would be of lesser amplitude than the dependence of T_a on D_c/V_{pl}, as discussed in section 4.5.4.

With a mechanism of nucleation so different between the finite fault and the SRM, it is actually interesting that the qualitative predictions for the seismicity rate of the SRM (equation (4.19)) reproduce the seismicity rate of the finite-fault model so well (Figure 4.11).

4.7.2 High sensitivity of finite faults to stress perturbations

The fact that the finite fault is more sensitive to step-like stress perturbations than the predictions of the SRM had already been pointed out by *Kaneko and Lapusta* (2008). They showed that nucleation zones that contribute to aftershock response do not always satisfy the approximation made by *Dieterich* (1994) that $V\theta/D_c \gg 1$ when the perturbation is applied. Replacing this approximation by the condition $V\theta/D_c = 1$ before the perturbation and $V\theta/D_c \gg 1$ only after the perturbation, they proposed a new expression for the amplitude of the response of the seismicity to a step-like stress perturbation (equation (D11) in *Kaneko and Lapusta* (2008)), which in our case would give $\beta_{\mathrm{KL}} \approx 1.6$. This estimate is therefore slightly closer than β_{SRM} to the values obtained in Figure 4.17 ($\beta = 2.6$ for $V_{pl} = 0.1$ cm/yr and $\beta = 3.7$ for $V_{pl} = 1$ and 10 cm/yr), but still underestimates the amplitude of the response. Applying the approach of *Kaneko and Lapusta* (2008) to the case of a harmonic stress perturbation leads to a response of amplitude $\beta_{\mathrm{KL}} = \Delta\tau/(b-a)\sigma$, twice as large as the predictions of the SRM since in our simulations a is twice as large as $b - a$. These modifications to the SRM could partly explain the larger sensitivity of the finite fault, since at periods $T \ll T_a$, $\beta/\beta_{\text{SRM}} \approx 3$ (Figure 4.12), but the finite fault is still more sensitive than this modification would prescribe.

Besides, the considerations by *Kaneko and Lapusta* (2008) do not capture the period-dependent response of the finite fault at periods $T < T_a$. *Ader et al.* (2012b) showed that the amplitude of slip-rate variations of a spring-slider system with rate-strengthening rheology under harmonic stress perturbations may be period dependent at periods $T < T_a$. In their study, the period dependence comes from the fact that if the period T of the perturbation is smaller than the period $T_\theta = 2\pi\theta_{\text{ss}} = 2\pi D_c/V_{\text{ss}}$, where θ_{ss} is the the characteristic time of evolution of the state variable (equal to the steady-state value of the state variable), and V_{ss} is the steady-state creep rate, the amplitude of the variations of creep rate is $\Delta V/V_{\text{ss}} = \Delta\tau/a\sigma$, whereas when $T > T_\theta$, $\Delta V/V_{\text{ss}} = \Delta\tau/(a-b)\sigma$. Such a period-dependent response may therefore be at play in the rate-strengthening part of the finite fault, but it does not appear be responsible for the period-dependent response of the seismogenic patch. Supplementary Figure S9 shows the response of the finite fault where we have imposed $b = 0$ in the rate-strengthening areas of the fault, so that the response of the rate-strengthening areas cannot be period dependent any more since $a - b = a$. Such a modification does not alter the period-dependent response of the seismicity.

Basing their study on the behavior of a spring-and-slider system with rate-weakening rheology, *Perfettini et al.* (2001) demonstrated the possibility of a resonance of the slip rate in the nucleation zone with a harmonic perturbation of normal or shear stress at the critical period $T_c = 2\pi\sqrt{a/(b-a)}(D_c/V_N)$, where V_N is the loading velocity on the spring-slider. However, that resonance would have a noticeable effect only for perturbing periods T within 5% of T_c. In the finite-fault simulations presented earlier, the increase of correlation is noticeable at periods two orders of magnitude around the critical period T_a (Figure 4.12). Since the fault parameters a, b and D_c are constant in the simulations, this would require the background slip velocity V_N in the nucleation zone to be such that $T_a \approx T_c$, and vary over approximately two orders of magnitude. Supplementary Figure S3 shows that the velocity in the nucleation zone is of the order of 0.1 cm/yr during most of the nucleation phase, and ranges from 0.04 cm/yr to 0.2 cm/yr at the onset of nucleation, corresponding to resonance periods T_c ranging from 0.02 to 0.1 years. The amplitude of the response of the seismicity on the finite fault to harmonic shear-stress perturbations is at least 3 times larger than the predictions of the SRM for periods ranging from 10^{-3} years to 1 year, therefore larger than the range possibly covered by T_c. Therefore, this resonance phenomenon probably explains a part of the high correlation around these periods, but cannot alone explain the high sensitivity of the finite fault to stress perturbations.

We showed earlier that events correlate with the shear-stress perturbation independently of their magnitude (Figures 4.7 and 4.12). An event nucleates on the fault when the nucleation zone (i.e., the zone at the edge of but within the seismogenic patch that creeps) reaches a critical size, and what then determines the size of the event is the distribution of pre-stresses on the fault (e.g., *Kaneko and Lapusta*, 2008). The fact that there is no obvious relation between the size of an event and its correlation with the stress perturbation therefore indicates that the correlation with the loading stress happens at the nucleation of events. Since there is no gradually growing nucleation zone during the interseismic period in the SRM, where the slider moves uniformly, this might explain the higher sensitivity of the finite-fault model to stress perturbations.

Under the premise that the correlation of the timing of events with the stress perturbation is determined at the nucleation, the high sensitivity of the finite fault can be heuristically understood following Griffith's theory (*Griffith*, 1924). Calling l the size of the nucleation zone, the energy δE_s necessary to increase this size by a length δl is the surface energy associated with the crack faces, i.e., the energy required to break the bounds holding both sides of the fault together. This surface energy is $\delta E_s = 2\gamma\delta l$, where γ is the surface free energy (e.g., *Segall*, 2010), depending on the material properties. Increasing the size of the nucleation zone by a length δl also releases the elastic energy $\delta E_{el} \propto K^2 \delta l$ (e.g., *Segall*, 2010), where the stress intensity factor $K \propto \Delta\tau\sqrt{l}$, and where $\Delta\tau$ is the stress drop. $\delta E_{el}/\delta l$ may therefore be expressed as $\delta E_{el} = \alpha l \Delta\tau^2 \delta l$, where α is another constant depending on the material properties. Therefore, the energy necessary in order to increase the size of the nucleation zone by a length δl is

$$\frac{\delta E}{\delta l} = 2\gamma - \alpha l \Delta\tau^2. \tag{4.27}$$

This energy thus decreases as the size of the nucleation zone increases. As long as the nucleation zone is small enough for $\delta E/\delta l$ to be positive, energy has to be brought to the fault in order to increase the size of the nucleation zone, which therefore grows slowly under the influence of the secular loading. Conversely, when the nucleation zone is large enough so that $\delta E/\delta l < 0$, increasing the size of the nucleation zone globally releases energy, so that the rupture propagates by itself and a seismic event is generated. At the onset of nucleation, $\delta E/\delta l \approx 0$. The stress drop can be expressed as $\Delta\tau = \tau(t) - \mu_d\sigma$, where μ_d is the dynamic friction on the fault and where the total shear on the fault $\tau(t)$ contains the external load perturbations, so that any small external perturbation of stress determines how the nucleation zone grows. The timing of the event will thus highly correlate with the perturbation.

Once an event has nucleated, it will rupture the entire seismogenic patch only if the distribution of pre-stresses on the seismogenic patch is adequate. Otherwise, the rupture will not be able to propagate and the event will only cover an edge of the patch. This would explain why we observed different magnitude distributions on the unperturbed and perturbed faults in the finite-fault model. If the perturbation modifies the relative evolutions of the size of the nucleation zone and of the pre-stresses on the fault, new magnitudes will be produced. For instance, whereas the finite fault does not produce any event of magnitude $M_{\text{lin}} < -0.2$ when left unperturbed (Figures 4.4b), it is able to produce events of magnitude as low as $M_{\text{lin}} = -0.6$ when perturbed (Figure 4.8 for harmonic and Figure 4.18 for step-like shear-stress perturbation). This indicates that in some cases, the perturbation causes the nucleation zone to grow and reach its critical size while the pre-stresses on the fault are still lower than they would be at this point on an unperturbed fault, which leads to smaller events. The fact that both $M_{\text{lin}} \sim 0.95$ and $M_{\text{lin}} \sim 1.2$ events disappear from the perturbed fault indicates that in some cases, the perturbation either delays or accelerates the nucleation when an unperturbed fault would have nucleated. This indeed indicates a high sensitivity of the propagation of the crack tip to any stress perturbation at the onset of failure. We also pointed out the occurrence of a larger proportion of $M_{\text{lin}} \sim 1.2$ events following a stress step than on an unperturbed fault. These events are most likely events that are already close to nucleating when the stress step is applied: the stress step therefore increases the already-high pre-stresses on the fault without changing much the timing of the event, yielding events rupturing the entire patch with a larger-than-average slip.

The variations of pre-stresses due to applied stress perturbations on the fault therefore appear to have an impact on the size of the events produced, whereas the timing of events is rather related to the modified growth of the nucleation zone under the influence of the stress perturbation. Understanding how the nucleation zone grows under the influence of a stress perturbation is therefore most likely a prerequisite for the establishment of a fully analytical framework describing the correlation of the timing of events with a given stress history in terms of the problem's parameters.

4.7.3 Implications for the estimation of $a\sigma$

A recurring feature in the behavior of a finite fault undergoing a stress perturbation is that the amplitude of the response is much larger than the predictions of the SRM. This observation calls for caution regarding the estimates of $a\sigma$ computed from the response of seismicity to stress perturbations, which are usually based on the predictions of the SRM (*Cochran et al.*, 2004; *Bettinelli et al.*, 2008; *Gross and Kisslinger*, 1997; *Gross and Bürgmann*, 1998; *Toda et al.*, 1998, 2012).

In order to compute $a\sigma$, one generally estimates the amplitude β of the relative variations of seismicity rate in response to a stress perturbation. In the case of periodic variations, β is half the peak to peak value of the seismicity rate, whereas for an aftershock sequence, a fit of equation (4.23) to the cumulative number of events is necessary to estimate β. Considering that this amplitude can be expressed as a function of the problem parameters with the predictions of the SRM, $a\sigma = \Delta\tau/\beta$, where the amplitude $\Delta\tau$ of the perturbation can be evaluated provided the fault geometry at the location of the events considered.

However, the finite fault simulations indicate that the amplitude of the correlation is much larger than the predictions of the SRM, so that it may be possible to observe a strong response of the seismicity to a shear stress perturbation without necessarily requiring very small values of $a\sigma$.

In the case where the finite fault is subjected to harmonic stress perturbations, the amplitude of the response at $T = 0.0027$ years is 5 times larger than what would be predicted by the SRM. This cannot be due to a resonance with the return period of events on the fault, as $T = 0.0027$ years is two orders of magnitude less than the natural periodicities on the fault (Figure 4.5). At $T = 0.027$ years, the amplitude of the response is about 20 times larger than the predictions of the SRM. As a result, using the SRM to explain the large amplitude of the seismicity rate variations at $T = 0.027$ years, knowing that the perturbation has a $\Delta\tau = 3$ kPa amplitude yields $a\sigma \approx 2$ kPa, where in fact $a\sigma = 40$ kPa. This could have important implications for values of $a\sigma$ that have been estimated on real faults from their response to a periodic perturbation (*Cochran et al.*, 2004; *Bettinelli et al.*, 2008). The actual value of $a\sigma$ in the seismogenic zone might actually be orders of magnitudes larger than the value computed with the SRM, which would then allow the effective normal stress to be as large as the overburden minus hydrostatic pore pressure at seismogenic depths.

The same conclusions remain true when estimating $a\sigma$ from aftershock sequences with the SRM predictions (*Gross and Kisslinger*, 1997; *Gross and Bürgmann*, 1998; *Toda et al.*, 1998, 2012). *Gross and Kisslinger* (1997) and *Gross and Bürgmann* (1998) actually used the total number of aftershocks N_a to first estimate $\dot{\tau_a}$ from the predictions of the SRM with equation (4.25) and then to obtain t_a from the fit of equation (4.23) to the cumulative number of events, to finally estimate $a\sigma = t_a\dot{\tau_a}$. This procedure is equivalent to estimating $a\sigma = \Delta\tau/\beta$ from the β-value from the fit. As our finite-fault simulations show, the observed value of β is greater than the predictions from the SRM, which would lead to underestimating $a\sigma$, as in the case of the harmonic response.

Similar issues with estimating $a\sigma$ from observations using the spring-slider-based SRM equations have been pointed out by *Kaneko and Lapusta* (2008) for a different finite-fault problem.

4.8 Conclusions

The response of a continuum model of finite seismogenic patch with rate-and-state rheology surrounded by creeping areas subjected to a stress perturbation appears to mimic the observed behavior of seismogenic faults in lab experiments and in nature. To some extent, the results can be qualitatively reproduced by Spring-slider Rate-and-state Models (SRMs) such as the model proposed by *Dieterich* (1994), although some of the features observed in the lab or on natural faults can only be successfully reproduced by the finite-fault model. Quantitatively, the finite fault is much more sensitive to stress perturbations than the SRM, suggesting that using the results of the SRM to infer fault properties from observations might lead to biased estimates.

Proceeding by analogies between the SRM and the finite-fault model, a few features observed on the finite fault are captured by the SRM. The first one is the qualitative shape of the seismicity rate in response to either a harmonic stress perturbation or a stress step, which can be fit by the predictions of the SRM (Figure 4.11 for the response to a harmonic perturbation and Figure 4.17 for the response to a stress step). In the case of the harmonic response, besides, the seismicity rate is approximately in phase with the stress perturbation at periods shorter than a critical period T_a, and with the stress rate at periods larger than T_a, in compliance with the predictions of the SRM (Figure 4.2). Moreover, both the characteristic period T_a dividing between these two regimes and the characteristic duration of the response to a stress step t_a seem to be inversely proportional to the loading rate on the patch, similar to their relation in the SRM: $T_a = 2\pi t_a = 2\pi a\sigma/\dot{\tau_a}$.

However, the paramount difference with the SRM is quantitative: the amplitude of the response to stress perturbations is much larger on the finite fault, sometimes by several orders of magnitudes. If one were to interpret the output of the finite-fault simulations with the formulae yielded by the SRM in order to compute fault properties of the seismogenic patch, one would systematically underestimate the product $a\sigma$, sometimes by a few orders of magnitude. This calls for caution when using these same formulae to estimate fault properties on real faults from the seismicity variations ensuing stress perturbations. In addition, where the SRM was unable to explain any increase of the amplitude of the response to harmonic stress perturbations with the perturbing period, yet observed in lab experiments (*Lockner and Beeler*, 1999; *Beeler and Lockner*, 2003) and on the Nepalese seismicity (*Bettinelli et al.*, 2008; *Ader and Avouac*, 2013), the finite fault reproduces it for perturbing periods up to T_a. This increase of amplitude with perturbing period seems to go hand in hand with a phase lag between the seismicity and the stress perturbation up to $\pi/4$, a feature also unexplained by the SRM.

These differences are to be somehow expected, owing to some prominent disparities between both models. Conceptually, the main difference between the SRM and the finite-fault model is the fact that on the finite fault, spatial variations of slip are permitted along the fault. This enables the advent a growing nucleation zone within the rate-weakening patch, i.e., a zone at the edge of the seismogenic patch undergoing creep, which does not exist in the SRM. If the notion of a critical nucleation size does exist in the SRM, no actual evolving nucleation zone is physically modeled. In the SRM, the slider moves as a block and the slip is therefore uniform in space. In the finite-fault model, the rupture initiates when the nucleation zone reaches a critical size, and the distribution of pre-stresses on the fault determines how far the rupture propagates, i.e., the size of the event. The timing of the event is thus directly linked to the growth process of the nucleation zone. What determines how an applied perturbation will affect the timing of events is thus how it will affect the evolution of the nucleation zone at the onset of nucleation, when the growth of the nucleation zone precisely becomes the most sensitive to any stress perturbation. This high sensitivity of the growth of the nucleation zone at the onset of nucleation most likely explains why the finite fault is overall much more sensitive than the SRM to stress perturbations. Besides, a stress perturbation modifies the size of the nucleation zone without necessarily bringing the pre-stresses to the corresponding configuration they would have had on an unperturbed fault. This might result in configurations of pre-stresses at the onset of nucleation on the fault subjected to a stress perturbation that would have never prevailed at the onset of nucleation on an unperturbed fault. This justifies why the stress perturbation appears to modify the distribution of magnitudes of the events produced on the finite fault, instead of simply modifying the timing of events as is assumed in the SRM.

The similarities between the SRM and the finite-fault model may be exploited in order to better understand and quantify the behavior of the finite fault. Notably, both models exhibit a characteristic time t_a for the response of the seismicity to a stress perturbation. In the SRM, the characteristic time of relaxation of the seismicity rate in response to a stress step (characteristic duration of aftershock sequence) is $t_a = a\sigma/\dot{\tau}_a$, where the background stressing rate $\dot{\tau}_a = kV_{pl}$ is well defined, k being the spring stiffness and V_{pl} the loading plate velocity. For the response to harmonic stress perturbations, $T_a = 2\pi t_a$ is the characteristic period separating between the regime where the seismicity rate correlates with the stress perturbation (periods $T \ll T_a$) from the one where it correlates with the stress-rate perturbation (periods $T \gg T_a$). These notions of a characteristic time t_a and a characteristic period T_a also seem to appear in the finite-fault model, but these quantities are more complicated to interpret in terms of the physical parameters of the problem, since $\dot{\tau}_a$ varies in time and space on the fault during the interseismic phase. Besides, the relation $T_a = 2\pi t_a$ does not come up as clearly as for the SRM, an indication of the important nonlinearity of the behavior of the finite fault. It seems nevertheless possible to derive an expression of T_a for the finite fault inspired from the expression yielded by the SRM. Given that the response of the fault to a stress perturbation is determined by the sensitivity of the nucleation zone at the onset of nucleation, we take for the background stressing rate $\dot{\tau}_a$ the stress rate close to the crack tip at the onset of nucleation. This results in a characteristic period T_a described by equation (4.26), i.e., proportional to the ratio D_c/V_{pl}, and relatively independent on the normal stress σ. Testing the influence of these three parameters on T_a seems to support this result.

Unlike in the SRM, V_{pl} and D_c also seem to have an effect on the amplitude of the response. In the case of V_{pl}, the effect is the opposite for the responses to harmonic and step-like perturbations: increasing V_{pl} increases the amplitude of the response to a stress step whereas it decreases the amplitude of the response to a harmonic perturbation at the critical period T_a. This is another illustration of the nonlinearity of the response of the finite fault with the problem's parameters. In the case of harmonic perturbations, varying V_{pl} can modify the amplitude of the response by several orders of magnitudes. Similarly, increasing the characteristic distance D_c seems to make the finite fault slightly more sensitive to harmonic stress perturbations at the characteristic period T_a. The reasons for these dependences still remain unclear.

The influence of other parameters on the characteristic time t_a is more intricate to establish as it appears that their individual contributions cannot be singled out. The analysis of their influence requires additional work and is beyond the scope of the present study. The eventual goal of the presented simulations would be to be able to establish new laws describing the quantitative behavior of a rate-and-state fault in response to a harmonic or step stress perturbation, in order to be able to estimate fault parameters from observations of seismicity on natural faults, either from time variations of the seismicity rate in response to an oscillating stress perturbation, or from the time evolution of aftershock sequences.

Applying the results of these simulations to real cases of seismicity variations seems like an exciting and feasible project. The simulations presented here are conducted on an isolated seismogenic patch, with constant size and rheologic properties, and show that the seismicity on an unperturbed seismogenic patch seems to display its own characteristic periodicities. However, it is unlikely that observed periodic variations of seismicity in a catalog are due to the natural periodicities of the faults. Regional seismicity results from the combined seismicity of a population of seismogenic patches with different sizes and possibly, although less likely, slightly different fault properties. Therefore, even if all the seismogenic patches considered had similar characteristic periodicities, these periodicities would have no reason to be in phase with one another and would thus be unlikely to build up a coherent observable signal. Besides, the natural periodicity on a seismogenic patch corresponds mostly to the return period of events rupturing the whole patch, which directly depends on the patch size (the return period is proportional to the square root of the patch size on our 2D faults, see supplementary Figure S7). Seismicity observed in nature is most likely produced by a population of faults of a wide variety of sizes, and therefore with a wide variety of natural periodicities, so that the natural periodicities of each fault will get scrambled together and thus go undetected. Now, when a stress perturbation is applied on a population of faults of different sizes, the resulting variations of seismicity will remain coherent even if the rheologic properties of the patches (fault parameters a, b, D_c, loading velocity V_{pl}, etc.) vary slightly. Regionally, these properties are likely to undergo only small relative variations, so that the responses of the individual patches can coherently combine to build up a detectable signal. Observed variations of seismicity in nature are therefore most likely due to a forcing of some kind, and identifying first the forcing and then the relation between this forcing and the resulting variations of seismicity rate may provide precious elements about regional fault properties.

Thus, there is a real perspective to be able to decipher fault properties more accurately from the observation of the evolution of seismicity in response to stress perturbations from these simulations.

4.A Response of seismicity to harmonic stress perturbations in the Coulomb Failure Model

The CFM predicts that the seismicity rate on the fault should be proportional to the stressing rate. Calling r the seismicity rate on the fault under constant stressing rate $\dot{\tau_a}$, this implies that under a stress history $\tau(t)$, a fault would have the following seismicity rate:

$$\frac{R(t)}{r} = \frac{\dot{\tau}(t)}{\dot{\tau_a}}. \tag{4.A.1}$$

This, however, is only true if the stresses on the faults are distributed uniformly up to the Coulomb Failure Stress (hereafter CFS). Also, the seismicity rate can clearly not have negative values. Therefore, if $\tau(t)$ decreases, the seismicity will shut off and only resume when the stress on the fault will have increased back to a value greater than its past largest value. Otherwise, all the faults that have not ruptured yet will have stresses less than the CFS. Calling $\tau_f(t)$ the increasing envelope of $\tau(t)$ (Figure 4.A.1), the seismicity rate on the fault should actually be written

$$\frac{R(t)}{r} = \frac{\dot{\tau_f}(t)}{\dot{\tau_a}}. \tag{4.A.2}$$

These stresses as well as the corresponding seismicity rate are represented in Figure 4.A.1, in the case where the shear stress $\tau(t)$ decreases at some point.

For a stress history made of a component increasing at constant rate $\dot{\tau_a}$ and of harmonic variations of amplitude $\Delta\tau$ and period T

$$\tau(t) = \dot{\tau_a} t + \Delta\tau \cos\left(2\pi\frac{t}{T}\right), \tag{4.A.3}$$

the stress rate is increasing on the fault if $\dot{\tau_a} \geq 2\pi\Delta\tau/T$, i.e.

$$T \geq T_\tau = \frac{2\pi\Delta\tau}{\dot{\tau_a}}. \tag{4.A.4}$$

In this case, $\tau_f(t) = \tau(t)$ and the amplitude of the seismicity-rate variations is simply

$$\frac{\Delta R}{r} = \frac{T_\tau}{T}, \tag{4.A.5}$$

and so for a given perturbation amplitude, the seismicity-rate variations are inversely proportional to the period.

For small periods, such that $T \ll T_\tau$, $\tau_f(t) \neq \tau(t)$ and the seismicity rate is zero between

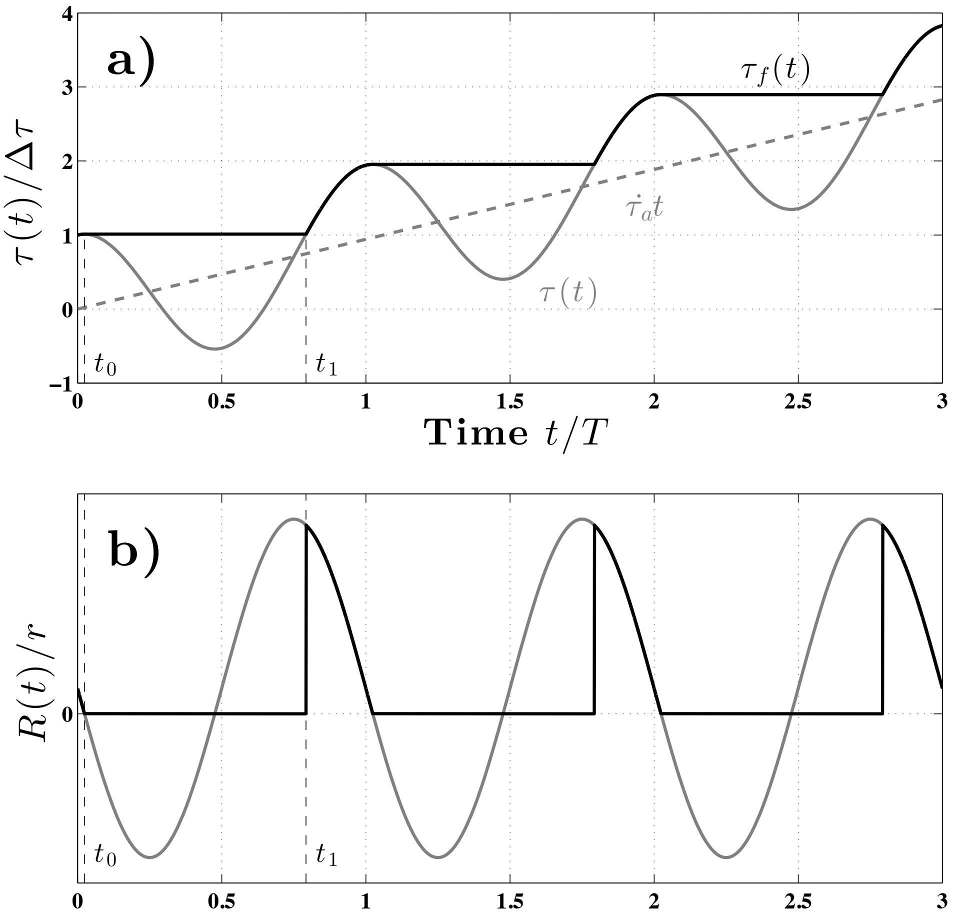


Figure 4.A.1: Evolution of the seismicity rate under a harmonic stress perturbation according to a Coulomb failure model (CFM). a) Shear stresses on the fault. Dashed grey line shows the secular loading $\dot{\tau_a}t$ on the fault, while the plain grey curve represents the total shear stress on the fault $\tau(t)$ (secular loading and harmonic variations). The black curve shows the stress actually "seen" by the seismicity, i.e., the increasing envelope of the stress on the fault. b) Seismicity rate on the fault corresponding to the shear stresses from upper plot. The grey curve shows the seismicity rate if it were truly proportional to the shear stress rate (and would thus be negative), while the black curve shows the actual (positive) seismicity rate on the fault. When the period of the stress becomes small enough, the actual amplitude of the seismicity-rate variations becomes smaller than the one of the stress rate (respectively black and grey curves on plot b)

times t_0 and t_1 within each period in Figure 4.A.1. It is maximum at time t_1, such that the amplitude of the variations is entirely defined by the value of the stress at t_1. The time t_0 is the first time within the period such that the stress rate cancels out. Taking the derivative of equation (4.A.3), this directly leads to, assuming $T \ll T_\tau$:

$$t_0 = \frac{T}{2\pi} \sin^{-1}\left(\frac{T}{T_\tau}\right) \approx \frac{T}{2\pi}\frac{T}{T_\tau}, \tag{4.A.6}$$

and

$$\tau(t_0) \approx \Delta\tau \left[1 + \frac{1}{2}\left(\frac{T}{T_\tau}\right)^2\right]. \tag{4.A.7}$$

The time t_1 is such that $t_1 > t_0$ and $\tau(t_1) = \tau(t_0)$. The approximation $T \ll T_\tau$ lets us write $t_1 = T - \varepsilon$, where ε is such that $\varepsilon/T \ll 1$ and it is solution of the equation:

$$\tau(t_1) \approx \dot{\tau_a} T \left(1 - \frac{\varepsilon}{T}\right) + \Delta\tau \left[1 - \frac{1}{2}\left(2\pi\frac{\varepsilon}{T}\right)^2\right]. \tag{4.A.8}$$

Equating equations (4.A.7) and (4.A.8) yields

$$\frac{\varepsilon}{T} = \sqrt{\frac{T}{\pi T_\tau}}. \tag{4.A.9}$$

The amplitude of the variations of seismicity rate is simply $\Delta R/r = R(t_1)/r - 1 = \dot{\tau}(t_1)/\dot{\tau_a} - 1$, which finally leads to the final expression:

$$\frac{\Delta R}{r} = 2\sqrt{\pi}\sqrt{\frac{T_\tau}{T}}. \tag{4.A.10}$$

The asymptotic lines for the Coulomb response in Figure 4.2 show that equations (4.A.5) and (4.A.10) give good approximations of the amplitude of the variations of seismicity rate in the cases where respectively $T/T_\tau \geq 1$ and $T/T_\tau \ll 1$.

4.B Response of seismicity to harmonic stress perturbations in the SRM

Based on a spring-slider system subjected to a rate-and-state friction law and a few other assumptions (see section 4.3), *Dieterich* (1994) proposed an equation to link the seismicity rate on a fault $R(t)$ to a given stress history $\tau(t)$. Assuming a constant normal stress on the fault, his equations (9) and (11) yield

$$\begin{cases} a\sigma\frac{d\gamma}{dt} = 1 - \gamma(t)\frac{d\tau}{dt}, \\ \frac{R(t)}{r} = \frac{1}{\dot{\tau_a}\gamma(t)}, \end{cases} \tag{4.B.1}$$

where r is the constant seismicity rate under constant stressing rate $\dot{\tau_a}$, a is a fault constitutive parameter and σ is the normal stress on the fault, assumed to be constant here. Taking the origin of time and shear stress so that $\tau(0) = 0$, and supposing that $R(t = 0^-) = r$, one can integrate the system of equations (4.B.1) to get

$$R(t) = \frac{re^{\tau(t)/a\sigma}}{1 + \frac{1}{t_a}\int_0^t e^{\tau(x)/a\sigma}\,dx}, \tag{4.B.2}$$

where $t_a = a\sigma/\dot{\tau_a}$ represents the characteristic time for the evolution of the seismicity rate.

Equation (4.B.2) can be easily linearized and solved in the case of small harmonic variations of the shear stress:

$$\tau(t) = \dot{\tau_a}t + \Delta\tau e^{i\omega t}, \tag{4.B.3}$$

where $\omega = 2\pi/T$, and we assume that $\Delta\tau \ll a\sigma$ (small perturbation assumption). In this case, once the steady-state regime is established (i.e., for times such that $t \gg t_a$) equation (4.B.2) reduces to

$$\frac{R(t)}{r} = \frac{1 + \frac{\Delta\tau}{a\sigma}e^{i\omega t}}{\frac{1}{t_a}\int_0^t e^{\frac{x-t}{t_a}}\left(1 + \frac{\Delta\tau}{a\sigma}e^{i\omega x}\right)\,dx}. \tag{4.B.4}$$

Simplifying equation (4.B.4), the seismicity rate can be written $R(t) = r + \Delta Re^{i\omega t}$ where the relative amplitude $\Delta R/r$ of the harmonic variations of seismicity rate verifies

$$\frac{\Delta R}{r} = \frac{\Delta\tau}{a\sigma}\frac{i\omega t_a}{1 + i\omega t_a}. \tag{4.B.5}$$

Equation (4.B.5) suggests two regimes of response of the seismicity to small stress perturba-

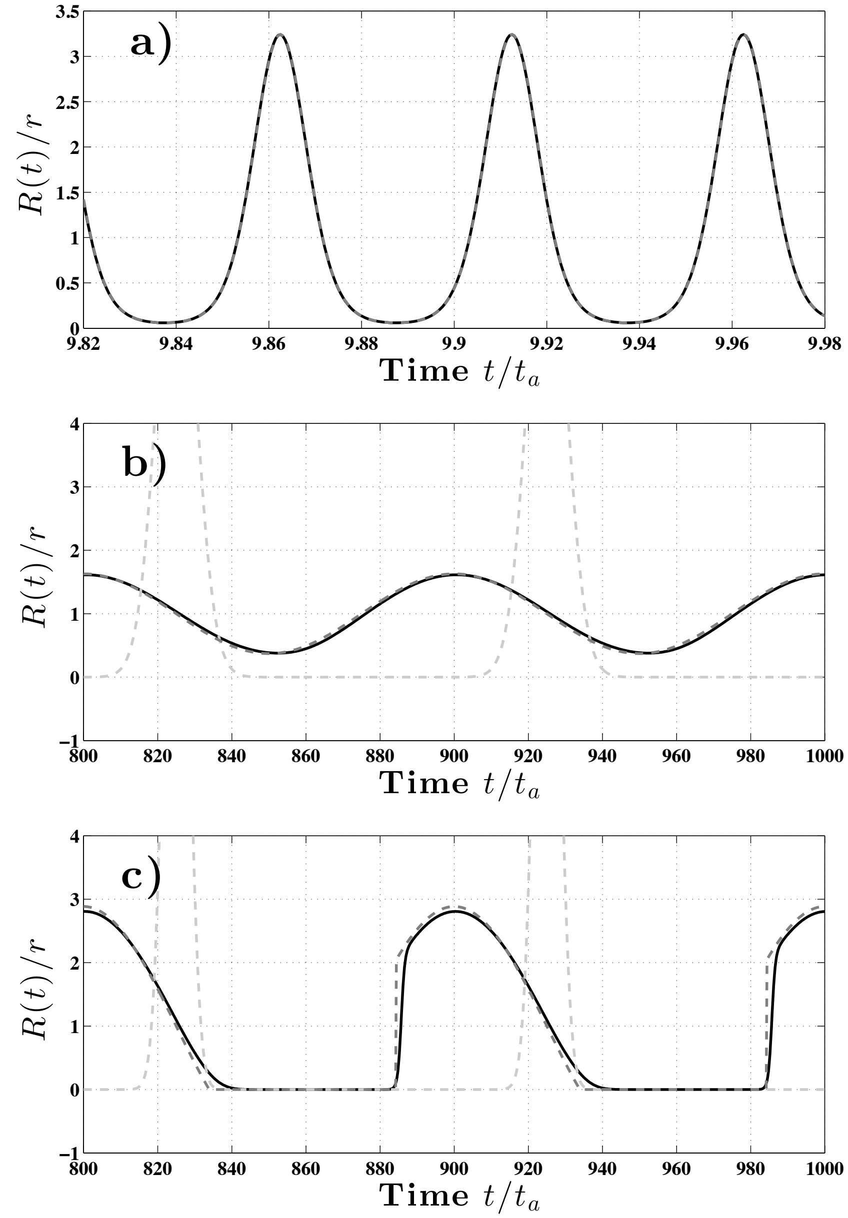


Figure 4.B.1: Evolution of the seismicity rate under a harmonic stress perturbation according to the SRM, simulated and analytical evolution. On all three plots, the plain black line shows the solution by solving equation (4.B.10) numerically. a) $\beta = 2$ and $T/t_a = 5 \times 10^{-2}$ (i.e., $T \ll t_a$ case), dashed grey line plots equation (4.B.13). b) $\beta = 10$ and $T/t_a = 100$ (i.e., $T \gg t_a$ case with $\beta \omega t_a \approx 0.6 < 1$), dashed dark-grey line plots equation (4.B.14) while dashed light-grey line plots equation (4.B.13) for comparison. c) $\beta = 30$ and $T/t_a = 100$ (i.e., $T \gg t_a$ case where $\beta \omega t_a \approx 1.9 > 1$), dashed dark-grey line plots equation (4.B.15) while dashed light-grey line plots equation (4.B.13) for comparison.

tions depending whether $\omega t_a \gg 1$ or $\omega t_a \ll 1$, i.e., depending on the value of the perturbing period T compared to the critical period

$$T_a = 2\pi t_a. \tag{4.B.6}$$

For perturbations at small periods ($\omega t_a \gg 1$, or $T \ll T_a$), the seismicity rate becomes proportional to the stress perturbation,

$$\frac{\Delta R}{r} = \frac{\Delta\tau}{a\sigma}, \tag{4.B.7}$$

and the response is thus period independent and is in phase with the shear stress. For perturbations of large period ($\omega t_a \ll 1$, or $T \gg T_a$), the seismicity rate follows the stress rate, as in the CFM:

$$\frac{\Delta R}{r} = \frac{i\omega\Delta\tau}{\dot{\tau_a}}. \tag{4.B.8}$$

This regime will be referred to as the Coulomb regime.

When the shear-stress perturbation becomes large, equation (4.B.2) can still be solved in the two configurations where $\omega t_a \ll 1$ and $\omega t_a \gg 1$. For a shear stress increasing at constant rate $\dot{\tau_a}$ with harmonic variations of amplitude $\Delta\tau$ and period T

$$\tau(t) = \dot{\tau_a} t + \Delta\tau \sin\omega t, \tag{4.B.9}$$

once the steady-state regime is established (i.e., for times such that $t \gg t_a$) equation (4.B.2) becomes

$$\frac{R(t)}{r} = \frac{e^{\frac{\Delta\tau}{a\sigma}\sin\omega t}}{\frac{1}{t_a}\int_0^t e^{\frac{x-t}{t_a}} e^{\frac{\Delta\tau}{a\sigma}\sin\omega x}\,dx}. \tag{4.B.10}$$

In the case of a harmonic perturbation at short periods $\omega t_a \gg 1$ ($T \ll T_a$), the periodic term in the integral in equation (4.B.10) evolves much faster than the exponential term, such that the denominator in equation (4.B.10) can be approximated by

$$\frac{1}{t_a}\int_0^t e^{\frac{x-t}{t_a}} e^{\frac{\Delta\tau}{a\sigma}\sin\omega x}\,dx \approx \left\langle e^{\frac{\Delta\tau}{a\sigma}\sin\omega t}\right\rangle \int_0^t \frac{1}{t_a} e^{\frac{x-t}{t_a}}\,dx \approx \left\langle e^{\frac{\Delta\tau}{a\sigma}\sin\omega t}\right\rangle, \tag{4.B.11}$$

where

$$\left\langle e^{\frac{\Delta\tau}{a\sigma}\sin\omega t}\right\rangle = \int_0^1 e^{\frac{\Delta\tau}{a\sigma}\sin 2\pi t}\,dt. \tag{4.B.12}$$

So in this case, the seismicity rate can simply be written

$$\frac{R(t)}{r} = \frac{e^{\frac{\Delta\tau}{a\sigma}\sin\omega t}}{\left\langle e^{\frac{\Delta\tau}{a\sigma}\sin\omega t}\right\rangle}. \tag{4.B.13}$$

Figure 4.B.1b shows that this expression provides a good approximation of the exact solution. Besides, one can easily verify that equation (4.B.7) is simply a linear approximation of this expression when $\frac{\Delta\tau}{a\sigma} \ll 1$.

In the opposite case where $\omega t_a \ll 1$ (i.e., $T \gg T_a$), the exponential term in the integral in equation (4.B.10) is significantly non zero only when the integration parameter x lies within a few t_a from t. Since $T_a \ll T$, the sine term does not vary much on this interval and can be approximated by $\sin\omega x \approx \sin\omega t + (x-t)\omega\cos\omega x$. With this approximation, and once the steady state regime has been reached (i.e., for times $t \ll t_a$), equation (4.B.10) reduces to:

$$\frac{R(t)}{r} = 1 + \frac{T_\tau}{T}\cos\omega t = \frac{\dot{\tau}(t)}{\dot{\tau}_a}, \tag{4.B.14}$$

which is analog to the Coulomb case. As in the Coulomb case, equation (4.B.14) is actually only valid in the case $T > T_\tau$ (i.e., $\frac{\Delta\tau}{a\sigma}\omega t_a < 1$), which is likely since $\omega t_a \ll 1$. However, strictly speaking, the ratio $\Delta\tau/a\sigma$ can be arbitrarily large, and thus T_τ might be greater than T. This case is similar to the case examined in appendix 4.A, and the result is the same: one simply has to replace $\tau(t)$ by its increasing envelope $\tau_f(t)$ to get the final result. Equation (4.B.14) can thus be directly replaced by

$$\frac{R(t)}{r} = \left(1 + \frac{T_\tau}{T}\cos\omega t\right)\mathcal{H}\left[\tau(t) - \tau(t_m(t))\right], \tag{4.B.15}$$

where $\mathcal{H}(t)$ is the Heavyside function ($\mathcal{H}(t) = 0$ for $t < 0$ and $\mathcal{H}(t) = 1$ for $t \geq 0$), and $t_m(t)$ is the time of previous shear stress maximum. Given the stress in equation (4.B.9), one has $t_m(t) = t - \mod (t - t_0, T)$, where t_0 is the time of first maximum stress within one period (same thing as in appendix 4.A), and is

$$t_0 = \frac{T}{2\pi}\cos^{-1}\left(-\frac{T}{T_\tau}\right). \tag{4.B.16}$$

Figure 4.B.1c shows that equation (4.B.15) provides a fairly good approximation of the exact solution.

Response of R&S faults to stress perturbations
***
Supplementary material

In order to test whether the distribution of times Δt between the instant where the stress step is applied and the last seismic event on the seismogenic patch can be described by an exponential PDF $p(\Delta t) = \frac{1}{\Lambda}\exp(-\Delta t/\Lambda)$, we follow the method described in the supplementary material of *Ader and Avouac* (2013). To each Δt we associate a Δt^u, which distribution is uniform if $p(\Delta t)$ is the right PDF for the distribution of the Δt:

$$\Delta t^u = \int_0^{\Delta t} p(t)\, dt. \tag{S1}$$

The Δt^u are by construction distributed over [0;1]. We then divide the [0;1] interval into b bins of equal width and compute the standard deviation σ_Λ of the normalized number of Δt^u falling within each bin. Given the number N_s of Δt^u, we know that σ_Λ should follow the distribution (*Ader and Avouac*, 2013):

$$p_s(\sigma_\Lambda) = \frac{2}{\Gamma\left(\frac{b}{2}\right)} \sqrt{\frac{N_s}{2}}^{\,b} \sigma_\Lambda^{b-1} e^{-\frac{N_s \sigma_\Lambda^2}{2}}, \tag{S2}$$

which has an expected value $\sqrt{b/N_s} \pm 1/\sqrt{2N_s}$. The computed value of σ_Λ for different values of Λ as well as the expected theoretical value with error bars are plotted in Figure S5, lower plot.

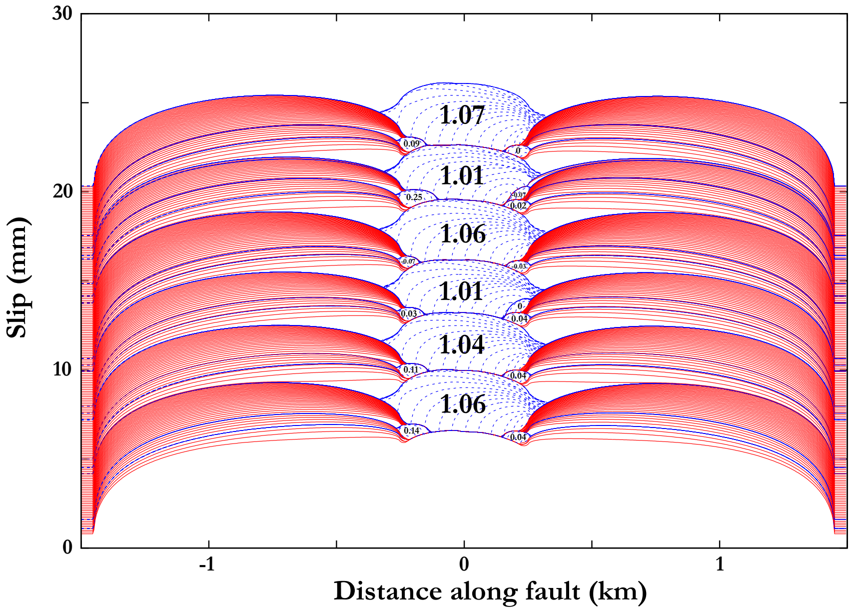


Figure S1: Displacement on the fault presented in section 4.4, when the fault undergoes a harmonic perturbation at period $T = 0.0027$ years. The red lines are plotted every 0.01 year during the interseismic period, while the blue dashed lines are plotted every 0.02 s when seismic rupture is occurring, i.e. when the maximum velocity on the fault is greater than 1cm/s. This pattern of slip is to compare to the slip on the unperturbed fault presented in Figure 4.4a. Once the fault is perturbed, the seismicity only consists of large $M_{\text{lin}} \approx 1.05$ events followed by two small ($-0.4 \leq M_{\text{lin}} \leq 0.4$) aftershocks, and sometime a small foreshock. The ratio of small to large events in this case is 2.25, indicating that on average there is a foreshock every 4 cycles. This well organized pattern of seismicity is very different from the seismicity on an unperturbed fault.

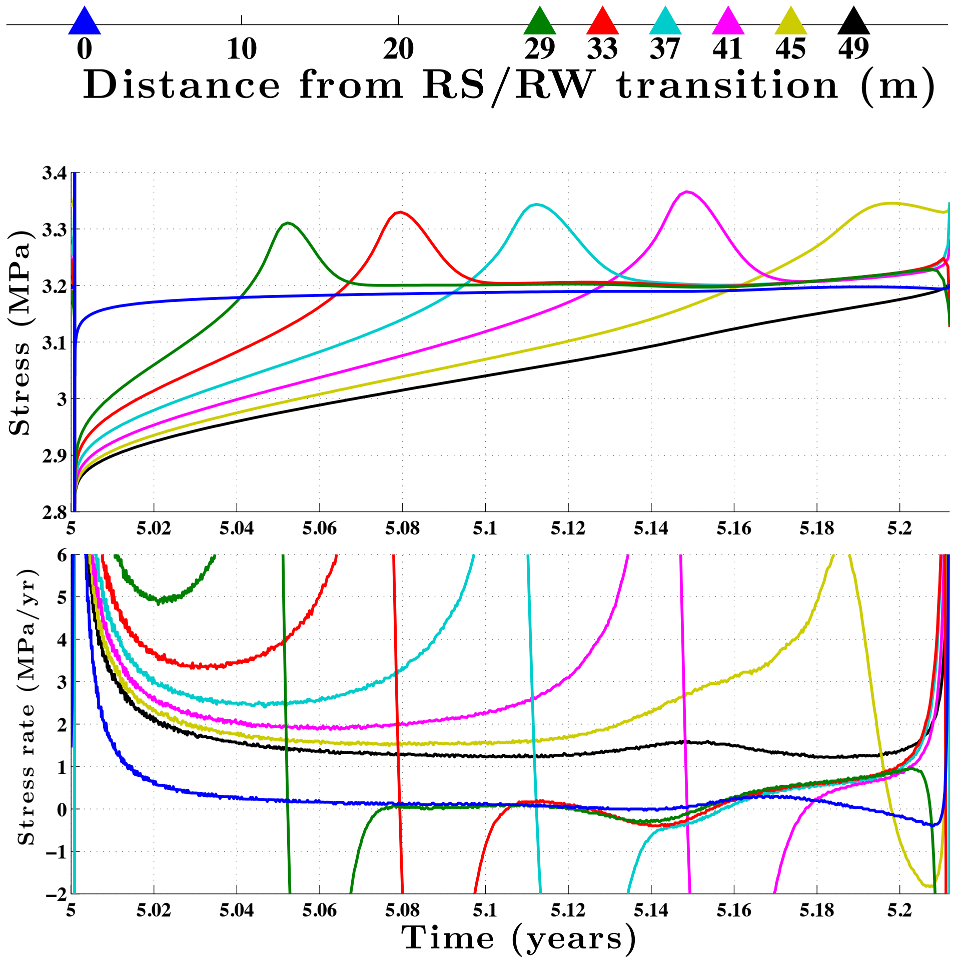


Figure S2: Evolution of the stress and of the stress rate with time in the nucleation zone. The shear stress and its time derivative are plotted at 7 points, one at the rate-weakening/rate-strengthening transition (blue curve), and the 6 other ones located inside the rate-weakening zone between 29 m and 49 m from the rate-strengthening/rate-weakening transition, equally spaced by 4 m from one another. The position of the points at which the stress is measured is indicated on the upper plot. At each of the points, the stress increases as the crack tip progresses toward the point of measurement, and then decreases to a common value once the point is within a creeping zone. The time at which the stress is maximum is the instant at which the crack tip is exactly at the point of measurement. The lower plot shows that the stress rate in the nucleation zone varies with time and localization within the nucleation zone, but that it is always larger than 1.25 MPa/yr (black line). The rupture nucleates when the crack tip is located between 45 and 49 m within the nucleation zone, which gives the critical nucleation size.

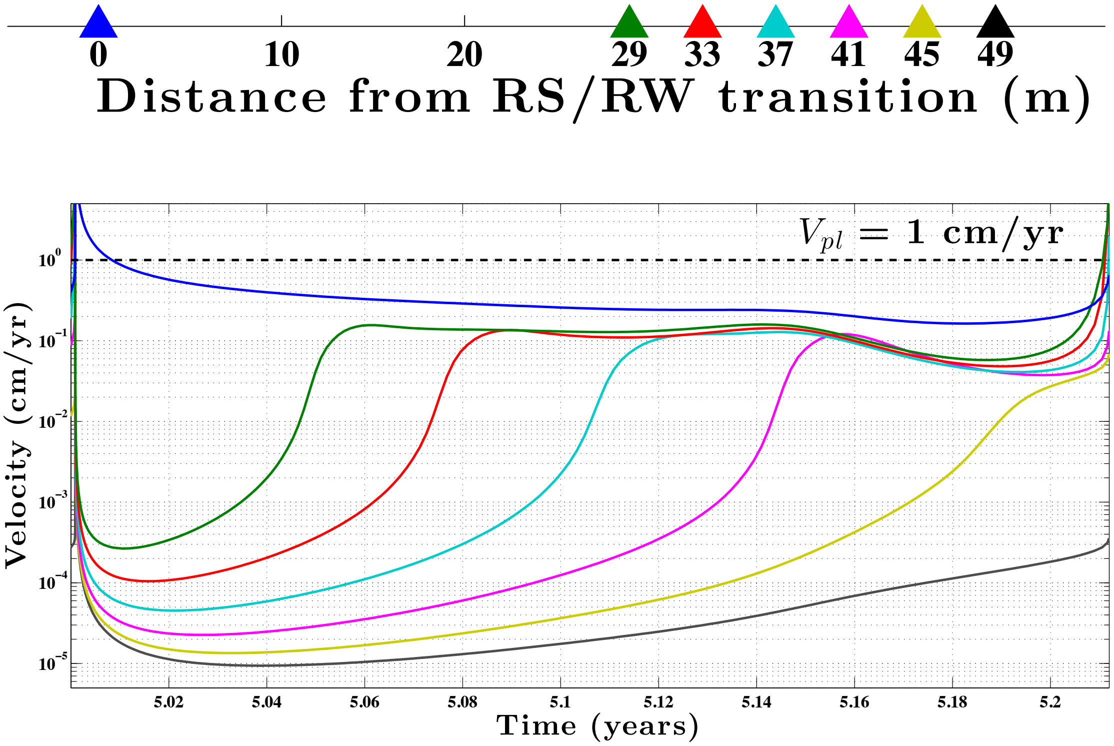


Figure S3: Evolution of the velocity with time in the nucleation zone, at the same locations as in Figure S2. The velocity in the nucleation zone is lower than the remote plate loading rate of $V_{pl} = 1$ cm/yr.

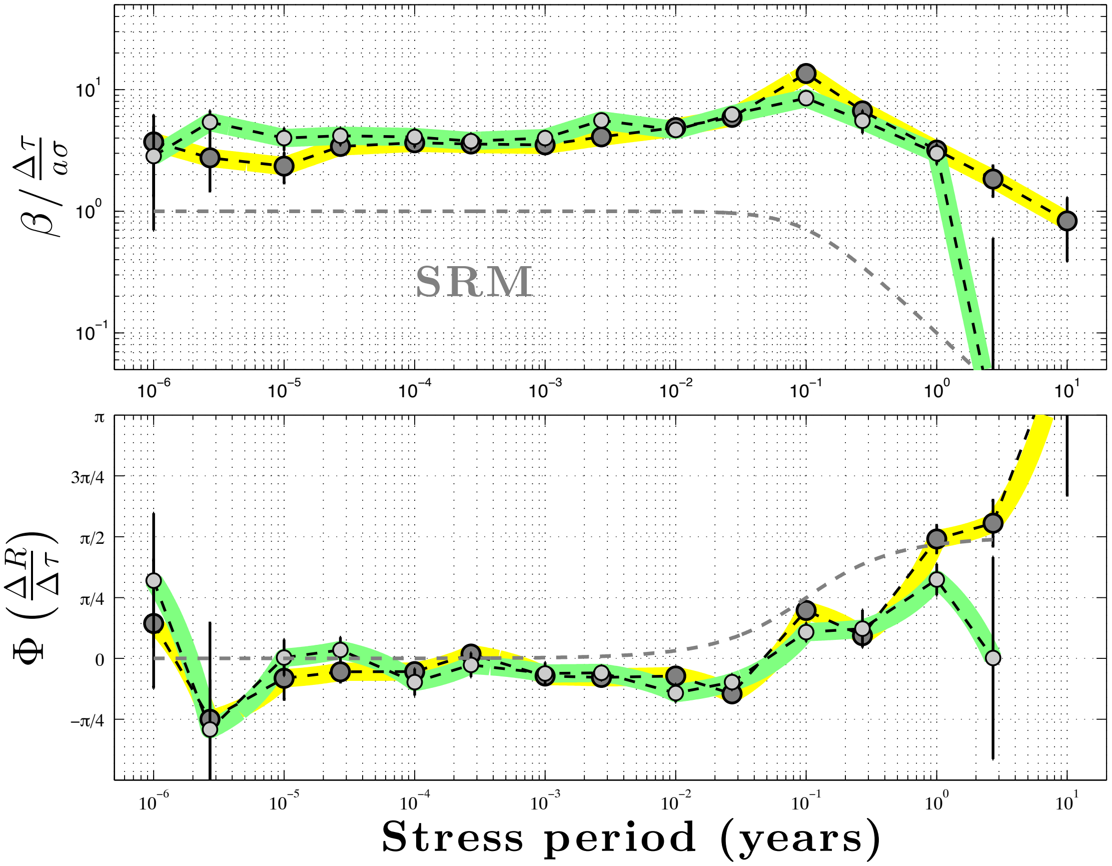


Figure S4: Comparison of the frequency response of faults with creeping zones of different sizes, for large events rupturing the entire seismogenic patch. The large dark grey circles overlying the yellow curve show the frequency response for the fault presented in section 4.4 which is the exact same plot as in Figure 4.12, and where the creeping zone is 1200m wide on each side of the seismogenic patch. The smaller light grey circles overlaying the green curve show the response for a fault with the same seismogenic patch, but where the creeping zone is only 450 m wide on each side. Both faults display the same response, indicating that as long as it is wide enough to avoid boundary effects, the actual size of the creeping zone does not have any significant impact on the frequency response of a seismogenic patch

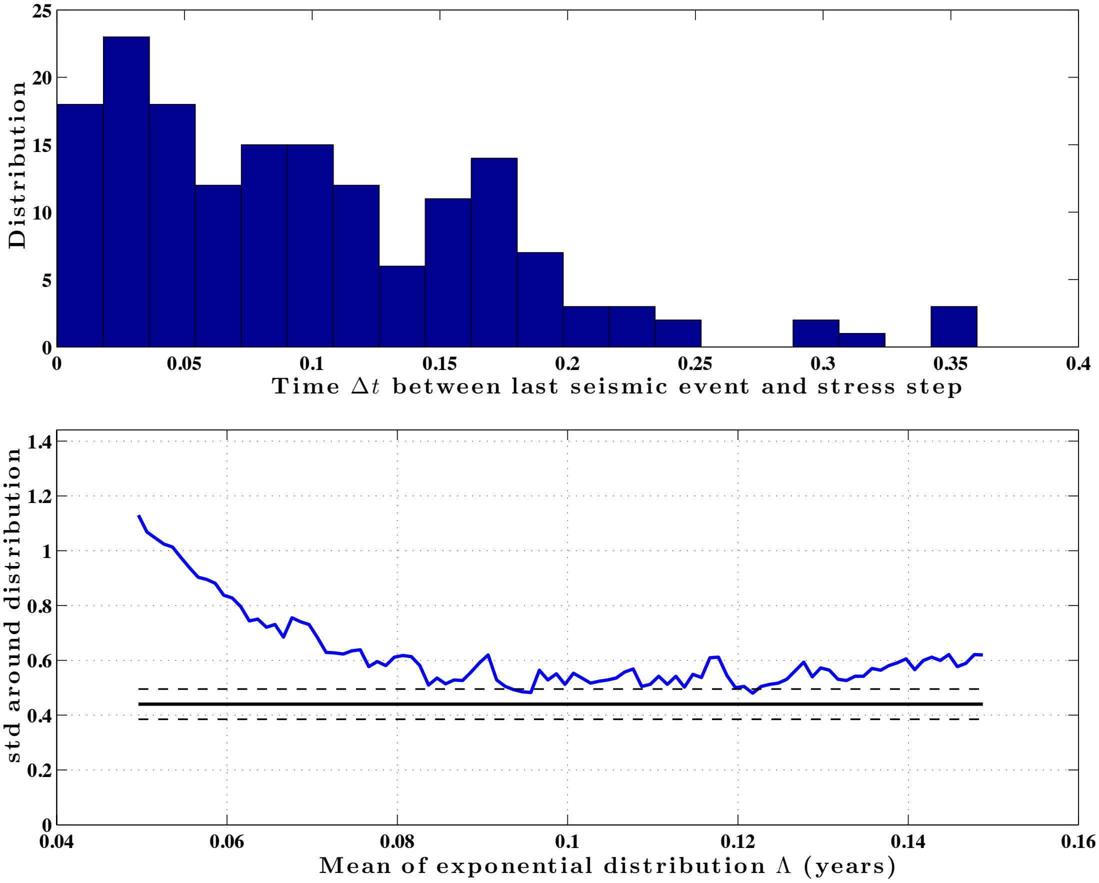


Figure S5: Upper plot: distribution of times between the last seismic event and the applied stress step. If the stress steps are applied at random times within the seismic cycle, the distribution of these inter times follows an exponential distribution, which mean is half the average inter event time. Lower plot: Test if the distribution of the Δt plotted in the upper plot can occur out of an exponential distribution, of a given mean. The blue line shows the standard deviation of the Δt redistributed along a uniform PDF if their initial distribution results from an exponential distribution of given mean, while the dark blue line indicates the expected the expected standard deviation, with 1-σ deviations indicated by the dashed lines. More details are given in the text.

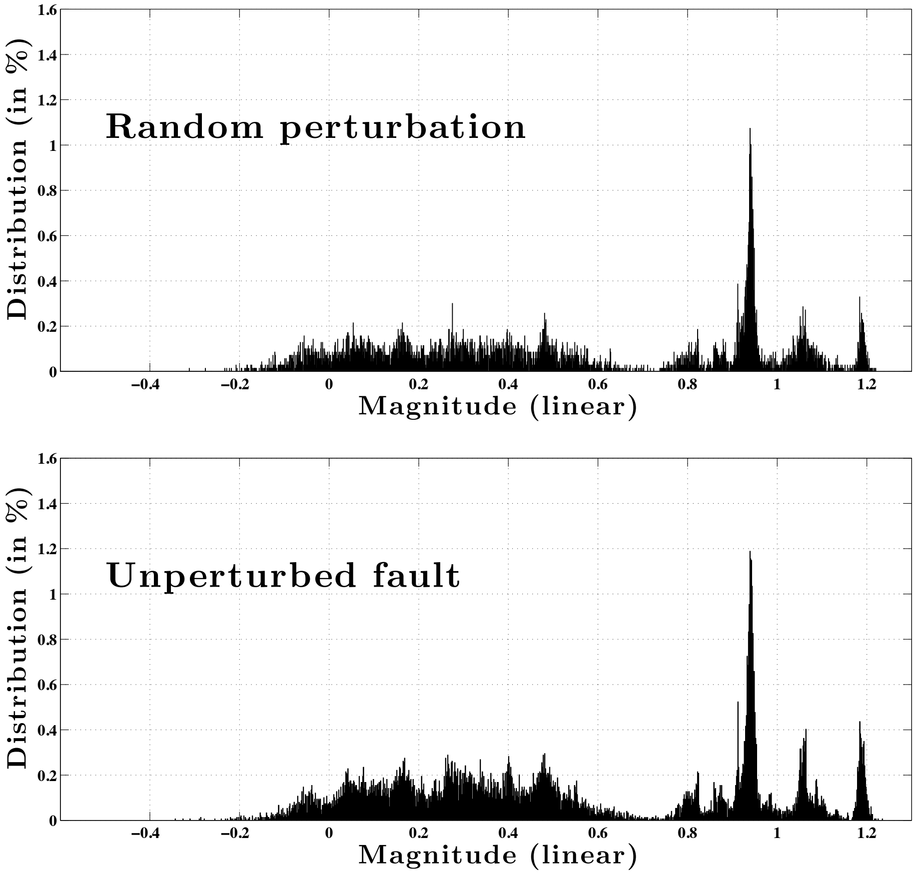


Figure S6: Comparison of the distribution of magnitudes of seismic events produced by the seismic patch on a fault undergoing a random perturbation in time. For comparison, the distribution of magnitudes of seismic events produced by the seismic patch on an unperturbed fault (figure 4.4b in the main paper) is showed on the lower plot, and shows that both distributions are the same. The random perturbation consists of steps of shear stress $\Delta\tau$ of amplitude uniformly distributed between -3 and 3kPa, applied at each time step during the simulation.

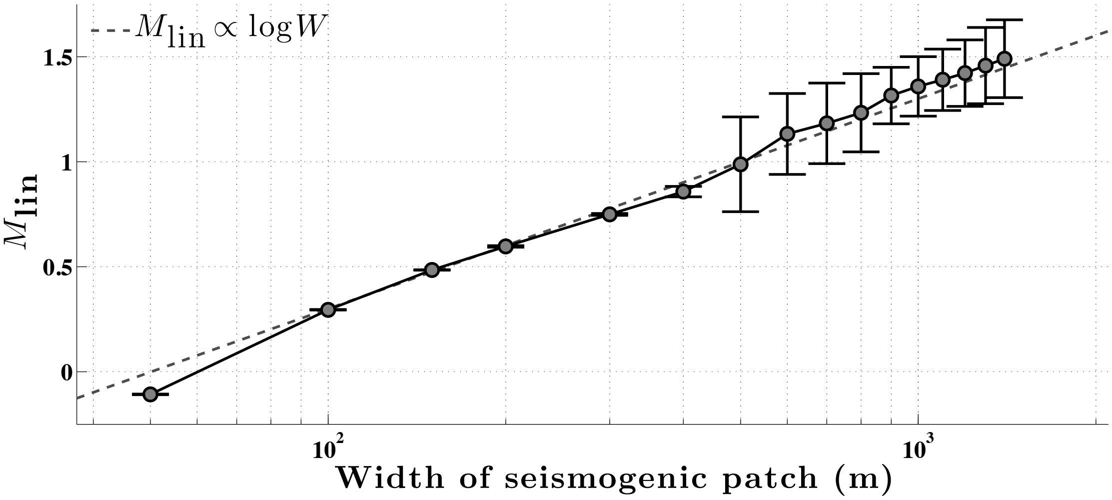


Figure S7: Distribution of linear magnitudes M_{lin} of seismic events that rupture the entire seismogenic patch, as a function of the width of the seismogenic patch. The error bars indicate the spread in magnitudes of events. For patches less than 400 m wide, the magnitudes of events are essentially always the same, while some complexity appears for patches wider than 500 m. This plots suggest that in first approximation, M_{lin} is directly proportional to the logarithm of the width of the seismogenic patch. Plugging this relation into equation (4.11) from the main paper leads to $\mathcal{M}_{\mathrm{lin}} \propto W^{3/2}$, which together with equation (4.15) indicates that the return period of large events rupturing the entire seismogenic patch depends on its width as $T \propto \sqrt{W}$.

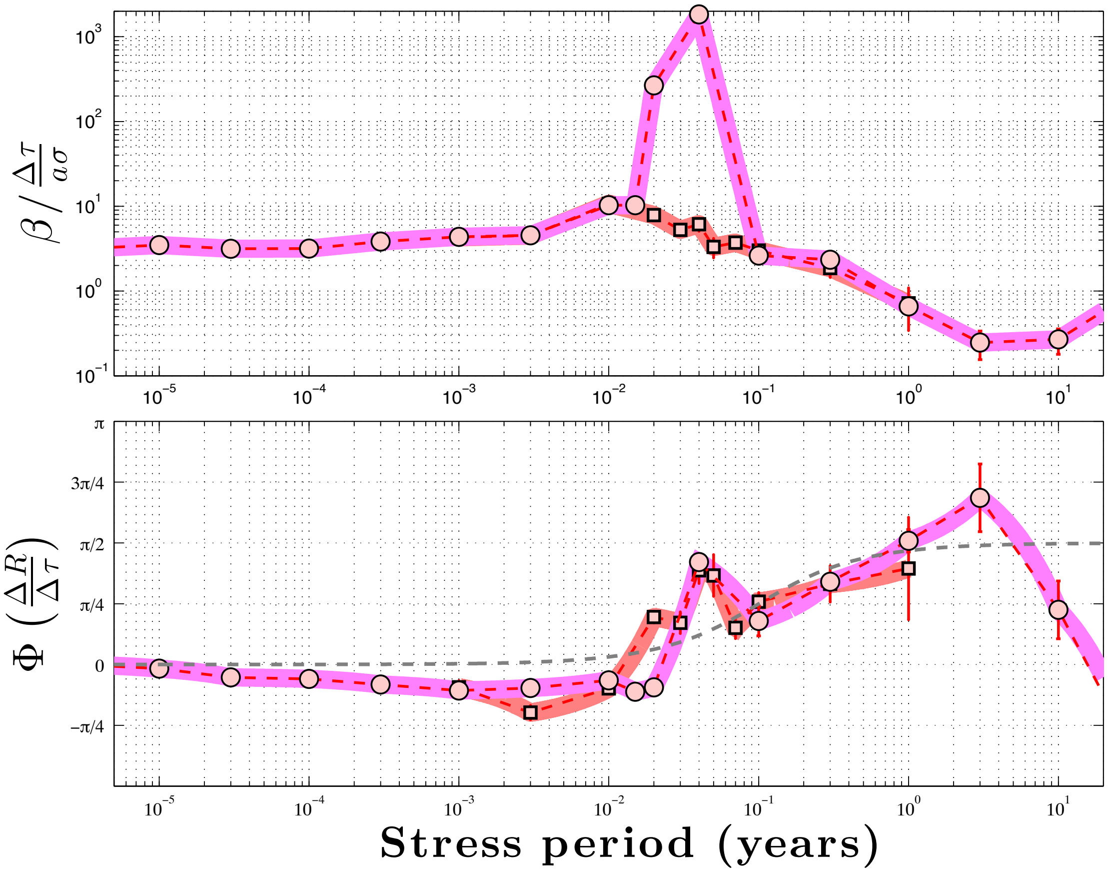


Figure S8: Plot showing the non-linearity due to resonances with the return period of characteristic events happening on the fault. In these two simulations, $D_c = 1$ μm and $\sigma = 1$ MPa, similarly to the red squares in Figure 4.15 from the main paper. In this Figure, only the amplitude of the stress perturbation $\Delta\tau$ is changed between the two curves: $\Delta\tau = 0.6$ kPa for the red squares (exactly same plot as red squares in Figure 4.15) while $\Delta\tau = 3$ kPa for the red circles. When $\Delta\tau = 3$ kPa, a resonance appears at $T \approx 0.05$ years, which is the return period of typical large events produced by the fault. This resonance can also be noted when $\Delta\tau = 0.6$ kPa but is of much less amplitude. When $\Delta\tau = 3$ kPa, this resonance impairs the determination of the critical period T_a, similarly to what happens at $D_c = 10$ μm in Figure 4.14 from the main paper (blue triangles).

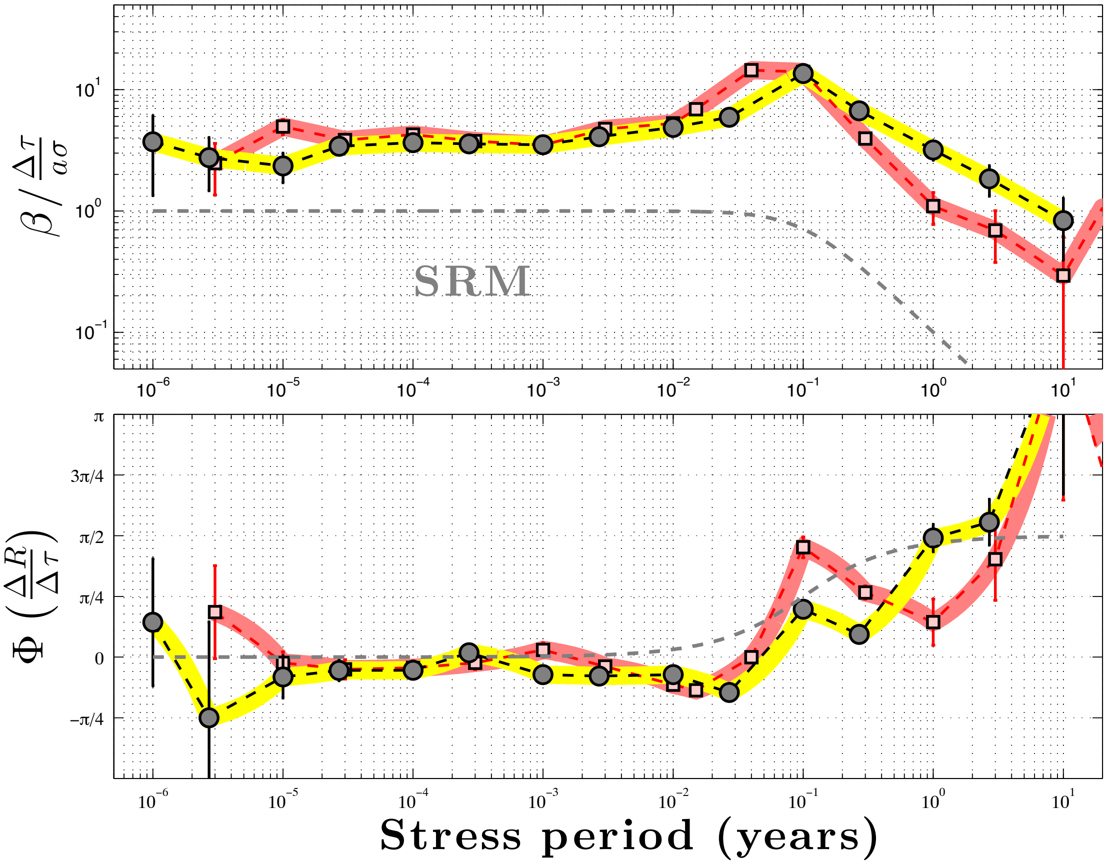


Figure S9: Plot showing that the increase of sensitivity of the seismogenic patch around period T_a is not due to the increase of the variations of slip rate in the creeping zone showed in *Ader et al.* (2012). The large dark grey circles overlying the yellow curve show the frequency response for the fault presented in section 4.4 which is the exact same curve as in Figure 4.12, while the pink squares overlaying the red curve show the response of a similar fault where fault parameter in the rate-strengthening zone is $b = 0$, such that $a - b = a$. As is showed in *Ader et al.* (2012), this would prevent any increase of amplitude of the slip rate variation in the creeping zone, and yet one can see an increase of the sensitivity of the seismicity.

Conclusion

The joint contributions of geodetic measurements, seismic observations and the development of numerical simulation, yield an extraordinary wealth of results and have the potential to bear much more. Geodetic observations in Nepal, today dominated by data from the GPS network, tomorrow maybe complemented by INSAR data, let us resolve the coupling pattern on the MHT and the convergence rate across the Nepal Himalaya, the two key ingredients to compute the rate at which moment deficit accumulates on the MHT underneath Nepal. The present day seismicity and major events of magnitude greater than 8 over the past 500 years are unable to balance this deficit, opening the gates to a fundamental question: How large of an earthquake can the MHT produce? Are the magnitudes 8 from the past as large of an event as the MHT can produce or rather the forewarning of a brewing gigantic Himalayan earthquake?

The answer to these questions requires a deepened understanding of fault properties and earthquake mechanics, and this is where numerical simulations take over. The observed response of the seismicity to stress perturbations bears the potential of providing us with the keys to unravel fault properties and understand earthquake mechanics. Numerical simulations provide a cracking instrument to analyze this response and dig out the relation between stress perturbations and associated variations of seismicity rate on a fault. The results of these simulations, examined in the present doctorate in the case of step-like and harmonic perturbations, the two configurations at play on the MHT in Nepal, look extremely promising in terms of deciphering fault parameters from observations of variations of seismicity on the MHT. They lay the bases for follow through work in order to build and hone the tools able to extract the fault properties on the MHT.

These fault properties on the MHT are a crucial piece of information in order to understand Nepalese seismicity, but the road towards a realistic simulation of the complete the seismic cycle in the Himalaya will have to go through the resolution of two other paramount computational challenges: the available computation power and taking into account the free surface of a non-vertical fault like the MHT in the numerical simulations. To this day, these two challenges still remain major impediments against the implementation of the precious numerical simulation that will hopefully determine wether or not the Himalaya can produce humongous earthquakes.

The available computational power is a problem because the locked part of the MHT is large, about 100 × 3000 km, while the smallest wavelength that has to be resolved in a dynamic rupture simulation is generally small, and does not scale with the total fault size. Instead, it is determined by the cohesion zone R_0 of the rupture front (*Lapusta and Liu*, 2009), which has been estimated by *Day et al.* (2005)

$$R_0 \approx \frac{G^* D_c}{b\sigma_{\text{eff}}},$$

where G^* is equal to the shear modulus G for mode III rupture and $G/(1-\nu)$ for mode II, with ν the Poisson's ratio, σ_{eff} is the effective normal stress, i.e., the lithostatic normal stress reduced by the pore pressure, and b and D_c are rate-and-state fault parameters. The size of this cohesion zone is thus inherent to the material considered and does not vary too much from one fault to another, and imposes a required maximum cell size for the simulation of the order of a hundred meters at the most. Given the current computational power available, such a small grid is manageable when simulating the seismic cycle on relatively small faults, such as Parkfield in California (*Barbot et al.*, 2012), which has dimensions of 15 × 30 kilometers, but remains way too intensive computationally for a fault an order of magnitude larger like the MHT.

Another limitation that was described earlier, is the boundary integral formulation of an algorithm like BICYCLE periodically stitches the fault to itself in order to deal with inertial effects. This is fine if one is looking to simulate the evolution of a fault which is self-similar at the edges, again like in the case of the seismogenic patch in Parkfield, which is entirely embedded within a creeping medium, but does not offer an suitable representation of the boundary conditions for a fault like the MHT, which is non vertical and has a free surface and is thus rate-weakening at the free edge and rate strengthening at the opposite edge at depth. A way to handle this free surface in the simulation yet remains to be found.

The Himalayan seismicity is an incredible mine of information, and its exploitation nurtures an exciting wealth of new challenges to overcome. In 2009, during my first field trip to Nepal in Lo-Mantang, in the remote Mustang area, a Buddhist monk explained to me that earthquakes were simply the jolts of the giant fish on whose back the world was built, whenever its guardian god released his attention. Our understanding of earthquakes has now come a long way from the fish model and places them out of the reach of some rancorous god trying to punish humans for their unpardonable sins. Earthquakes simply obey the laws of nature, the laws of physics, and like most other natural disasters there is no reason that the human mind will not find a way to crack open the secret of their prediction. The clock is ticking before the Himalaya produces its greatest earthquake, and we are still gathering the pieces needed to forecast it. We may have a few hundred years to solve the puzzle. Or maybe just a couple of days...

Bibliography

Ader, T., and J.-P. Avouac (2013), Detecting periodicities in earthquake catalogs using the Schuster test, application to Himalayan seismicity, *Earth and Planetary Science Letters, Submitted.*

Ader, T., J.-P. Avouac, J. Liu-Zeng, H. Lyon-Caen, L. Bollinger, J. Galetzka, J. Genrich, M. Thomas, K. Chanard, S. N. Sapkota, S. Rajaure, P. Shrestha, L. Ding, and M. Flouzat (2012a), Convergence rate across the Nepal Himalaya and interseismic coupling on the Main Himalayan Thrust: Implications for seismic hazard, *Journal of Geophysical Research - Solid Earth*, *117, B04403*, doi:10.1029/2011JB009071.

Ader, T., J.-P. Ampuero, and J.-P. Avouac (2012b), The role of velocity-neutral creep on the modulation of tectonic tremor activity by periodic loading, *Geophysical Research Letters*, *39, L16310*, doi:10.1029/2012GL052326.

Ader, T. J., N. Lapusta, and J.-P. Avouac (2013), Frequency response of rate-and-state seismogenic faults to shear stress perturbations, *Journal of Geophysical Research - Solid Earth*, *submitted.*

Altamimi, Z. (2009), The International Terrestrial Reference Frame (ITRF2005), *International Association of Geodesy Symposia*, *134, Part 2*, 81–82, doi:10.1007/978-3-642-00860-3.

Ambraseys, J. J., and J. Douglas (2004), Magnitude calibration of north Indian earthquakes, *Geophysical Journal International*, *159*, 165–206.

Armijo, R., P. Tapponnier, J. Mercier, and T. Han (1986), Quaternary Extension in Southern Tibet - Field Observations and Tectonic Implications, *Journal of Geophysical Research-Solid Earth and Planets*, *91*, 13,803–13,872.

Avouac, J.-P. (2003), Mountain building, erosion and the seismic cycle in the Nepal Himalaya, *Advanced Geophysics*, *46*, 1–79.

Avouac, J.-P. (2007), Dynamic Processes in Extensional and Compressional Settings - Mountain Building : From Earthquakes to Geological Deformation, *Treatise of Geophysics*, *6*, 377–439.

Avouac, J.-P., L. Bollinger, L. Lavé, R. Cattin, and M. Flouzat (2001), Le cycle sismique en Himalaya, *Compte Rendu de l'Académie des Sciences*, *333*, 513–529.

Banerjee, P., R. Bürgmann, B. Nagarajan, and E. Apel (2008), Intraplate deformation of the Indian subcontinent, *Geophysical Research Letters*, *35, L18301*, doi: 10.1029/2008GL035468.

Barbot, S., Y. Fialko, and Y. Bock (2009), Postseismic deformation due to the M_w 6.0 2004 Parkfield earthquake: Stress-driven creep on a fault with spatially variable rate-and-state friction parameters, *Journal of Geophysical Research*, *114, B07405*, doi: 10.1029/2008JB005748.

Barbot, S., N. Lapusta, and J.-P. Avouac (2012), Under the Hood of the Earthquake Machine: Toward Predictive Modeling of the Seismic Cycle, *Science*, *336*, 707–710.

Beeler, N., and D. Lockner (2003), Why earthquakes correlate weakly with the solid Earth tides: Effects of periodic stress on the rate and probability of earthquake occurrence, *Journal of Geophysical Research - Solid Earth*, *108*(B8), doi:10.1029/2001JB001518.

Beeler, N. M. (2009), Constructing Constitutive Relationships for Seismic and Aseismic Fault Slip, *Pure and Applied Geophysics*, *166*, 1775–1798.

Beeler, N. M., and D. A. Lockner (2003), Why earthquakes correlate weakly with the solid Earth tides: Effects of periodic stress on the rate and probability of earthquake occurrence, *Journal of Geophysical Research*, *108, NO. B8*, 2391, doi:10.1029/2001JB001518.

Bettinelli, P., J.-P. Avouac, M. Flouzat, F. Jouanne, L. Bollinger, P. Willis, and G. Chitrakar (2006), Plate motion of India and Interseismic strain in the Nepal Himalaya from GPS and DORIS measurements, *Journal of Geodesy*, *80*, 567–589.

Bettinelli, P., J.-P. Avouac, M. Flouzat, L. Bollinger, G. Ramillien, S. Rajaure, and S. Sapkota (2008), Seasonal variations of seismicity and geodetic strain in the Himalaya induced by surface hydrology, *Earth and Planetary Sciences Letters*, *266*, 332–344.

Bilham, R. (2004), Earthquakes in India and the Himalaya: tectonics, geodesy and history, *Annals of Geophysics*, *47(2)*, 839–858.

Bilham, R., P. Bodin, and M. Jackson (1995), Entertaining a great earthquake in Western Nepal: Historic activity and Geodetic test for the development of strain., *J. Nepal Geol. Soc.*, *11*, 73–88.

Bilham, R., K. Larson, J. Freymuller, and I. members (1997), GPS measurements of present-day convergence across the Nepal Himalaya, *Nature*, *386*, 1–94.

Blanpied, M. L., D. A. Lockner, and J. D. Byerlee (1995), Frictional slip of granite at hydrothermal conditions, *Journal of Geophysical Research*, *100*, 13,045–13,064.

Bollinger, L., J. Avouac, R. Cattin, and M. Pandey (2004), Stress buildup in the Himalaya, *Journal of Geophysical Research - Solid Earth*, *109, B11405*(B11), doi: 10.1029/2003JB002911.

Bollinger, L., F. Perrier, J.-P. Avouac, S. Sapkota, U. Gautam, and T. D.R. (2007), Seasonal modulation of seismicity in the Himalaya of Nepal, *Geophysical Research Letters*, *34, L08304*, doi:10.1029/2006GL029192.

Bosl, W. J., and A. Nur (2002), Aftershocks and pore fluid diffusion following the 1992 Landers earthquake, *Journal of Geophysical Research*, *107(B12)*, 2366, doi: 10.1029/2001JB000155.

Brooks, B. A., M. Bevis, K. Whipple, J. R. Arrowsmith, J. Foster, T. Zapata, E. Kendrick, E. Minaya, A. Echalar, M. Blanco, P. Euillades, M. Sandoval, and R. J. Smalley Jr (2011), Orogenic-wedge deformation and potential for great earthquakes in the central Andean backarc, *Letters to Nature*, *4*, 380–383.

Brown, J. R., G. C. Beroza, S. Ide, K. Ohta, D. R. Shelly, S. Y. Schwartz, W. Rabbel, M. Thorwart, and H. Kao (2009), Deep low-frequency earthquakes in tremor localize to the plate interface in multiple subduction zones, *Geophysical Research Letters*, *36, L19306*, doi:10.1029/2009GL040027.

Bürgmann, R., M. Kogan, G. Steblov, G.M.and Hilley, V. Levin, and E. Apel (2005), Interseismic coupling and asperity distribution along the Kamchatka subduction zone, *Journal of Geophysical Research-Solid Earth*, *110*.

Burtin, A., L. Bollinger, J. Vergne, R. Cattin, and J. L. Nábelek (2008), Spectral analysis of seismic noise induced by rivers: A new tool to monitor spatiotemporal changes in stream hydrodynamics, *Journal of Geophysical Research*, *113, B05301*, doi: 10.1029/2007JB005034.

Cappa, F., J. Rutqvist, and K. Yamamoto (2009), Modeling crustal deformation and rupture processes related to upwelling of deep CO(2)-rich fluids during the 1965-1967 Matsushiro earthquake swarm in Japan, *Journal of Geophysical Research - Solid Earth*, *114, B10304*, doi:10.1029/2009JB006398.

Cattin, R., and J. Avouac (2000), Modeling mountain building and the seismic cycle in the Himalaya of Nepal, *Journal of Geophysical Research - Solid Earth*, *105*, 13,389–13,407.

Chen, Q., J. T. Freymueller, Q. Wang, Y. Zhiqiang, C. Xu, and J. Liu (2004), A deforming block model for the present-day tectonics of Tibet, *Journal of Geophysical Research*, *109, B01403*, doi:10.1029/2002JB002151.

Chen, W.-P., and P. Molnar (1977), Seismic moments of major earthquakes and the average rate of slip in Central Asia, *Journal of Geophysical Research*, *82(20)*, 2945–2969.

Chlieh, M., J. Avouac, K. Sieh, D. Natawidjaja, and J. Galetzka (2008), Heterogeneous coupling of the Sumatran megathrust constrained by geodetic and paleogeodetic measurements, *Journal of Geophysical Research*, *113*, B05,305.

Christiansen, L., S. Hurwitz, M. Saar, S. Ingebritsen, and P. Hsieh (2005), Seasonal seismicity at western United States volcanic centers, *Earth and Planetary Science Letters*, *240*, 307–321.

Christiansen, L. B., S. Hurwitz, and S. E. Ingebritsen (2007), Annual modulation of seismicity along the San Andreas Fault near Parkfield, CA, *Geophysical Research Letters*, *34, L04306*, doi:10.1029/2006GL028634.

Cochran, E. S., J. E. Vidale, and S. Tanaka (2004), Earth Tides Can Trigger Shallow Thrust Fault Earthquakes, *Science*, *306*, 1164–1166.

Dahm, T., S. Hainzl, and T. Fischer (2010), Bidirectional and unidirectional fracture growth during hydrofracturing: Role of driving stress gradients, *Journal of Geophysical Research - Solid Earth*, *115, B12322*, doi:10.1029/2009JB006817.

Day, S. M., L. A. Dalguer, N. Lapusta, and Y. Liu (2005), Comparison of finite difference and boundary integral solutions to three-dimensional spontaneous rupture, *Journal of Geophysical Research*, *110, B12307*, doi:10.1029/2005JB003813.

Dieterich, J. H. (1978), Time-dependent friction and the mechanics of stick-slip, *Pure and Applied Geophysics*, *116*, 790–806.

Dieterich, J. H. (1979a), Modeling of rock friction, 1, Experimental results and constitutive equations, *Journal of Geophysical Research*, *84*, 2161–2168.

Dieterich, J. H. (1979b), Modeling of rock friction, 2, Simulation of preseismic slip, *Journal of Geophysical Research*, *84*, 2169–2175.

Dieterich, J. H. (1992), Earthquake nucleation on faults with rate and state-dependent friction, *Tectonophysics*, *211*, 115–134.

Dieterich, J. H. (1994), A Constitutive Law for Rate of Earthquakes Production and its application to Earthquake Clustering, *Journal of Geophysical Research*, *99*(B2), 2601–2618.

Dieterich, J. H., and B. D. Kilgore (1994), Direct observation of frictional contacts: New insights for state-dependent properties, *Pure and Applied Geophysics*, *143*, 283–302.

Estrin, Y., and Y. Bréchet (1996), On a Model of Frictional Sliding, *Pure and Applied Geophysics*, *147(4)*, 745–762.

Felzer, K. R., and E. E. Brodsky (2006), Decay of aftershock density with distance indicates triggering by dynamic stress, *Nature Geoscience*, *441*, 735 – 738.

Freed, A. M. (2007), Afterslip (and only afterslip) following the 2004 Parkfield, California, earthquake, *Geophys. Res. Lett.*, *34, L06312*, doi:10.1029/2006GL029155.

Freymueller, J., S. Cohen, and H. Fletcher (2000), Spatial variations in present-day deformation, Kenai Peninsula, Alaska, and their implications, *Journal of Geophysical Research*, *105*, 8079–8101.

Fukuda, J., K. M. Johnson, K. Larson, and S. Miyazaki (2009), Fault friction parameters inferred from the early stages of afterslip following the 2003 Tokachi-oki earthquake, *Journal of Geophysical Research*, *114, B04412*, doi:10.1029/208JB006166.

Gomberg, J., P. Bodin, and P. A. Reasenberg (2003), Observing earthquakes triggered in the near field by dynamic deformations, *Bulletin of the Seismological Society of America*, *93*, 118 – 138.

Griffith, A. A. (1924), The Theory of Rupture, *Proceedings, First International Congress of Applied Mechanics*, pp. 55–63.

Gross, S., and R. Bürgmann (1998), Rate and state of background stress estimated from the aftershocks of the 1989 Loma Prieta, California, earthquake, *Journal of Geophysical Research*, *103, NO. B3*, 4915–4927.

Gross, S., and C. Kisslinger (1997), Estimating tectonic stress rate and state with Landers aftershocks, *Journal of Geophysical Research*, *102, NO. B4*, 7603–7612.

Gutenberg, B., and C. Richter (1954), *Seismicity Of The Earth And Associated Phenomena*, Princeton University Press.

Hainzl, S., and T. Fischer (2002), Indications for a successively triggered rupture growth underlying the 2000 earthquake swarm in Vogtland/NW Bohemia, *Journal of Geophysical Research - Solid Earth*, *107, NO. B12, 2338*, doi:10.1029/2002JB001865.

Hawthorne, J. C., and A. M. Rubin (2010), Tidal modulation of slow slip in Cascadia, *Journal of Geophysical Research*, *115, B09406*, doi:10.1029/2010JB007502.

Hearn, E., R. Bürgmann, and R. Reilinger (2002), Dynamics of Izmit earthquake postseismic deformation and loading of the Duzce earthquake hypocenter, *Bulletin of the Seismological Society of America*, *92*, 172–193.

Heaton, T. H. (1975), Tidal Triggering of Earthquakes, *Geophysical Journal of the Royal Astronomical Society*, *43*, 307–326.

Heki, K. (2003), Snow load and seasonal variation of earthquake occurrence in Japan, *Earth and Planetary Science Letters*, *207*, 159–164.

Heki, K., S. Miyazaki, and H. Tsuji (1997), Silent fault slip following an interplate thrust earthquake at the Japan Trench, *Nature*, *386*, 595–598.

Herman, F., P. Copland, J.-P. Avouac, L. Bollinger, G. Mahéo, P. Le Fort, S. Rai, D. Foster, A. Pêcher, K. Stüwe, and P. Henry (2010), Exhumation, crustal deformation, and thermal structure of the Nepal Himalaya derived from the inversion of thermochronological and thermobarometric data and modeling of the topography, *Journal of Geophysical Research*, *115, B06407*, doi:10.1029/2008JB006126.

Hernandez, G. (1999), Time series, periodograms, and significance, *Journal of Geophysical Research*, *104(A5)*, 10,355–10,368.

Herring, T., R. King, and S. McClusky (2009), *GAMIT Reference Manual and GLOBK Reference Manual, Release 10.3*, Mass. Inst. Technol.

Hill, D. P. (1993), Seismicity remotly triggered by the magnitude 7.3 Landers, California, earthquake, *Science*, *260*, 1617–1623.

Hodges, K. V., C. Wobus, K. Ruhl, T. Schildgen, and K. Whipple (2004), Quaternary deformation, river steepening, and heavy precipitation at the front of the Higher Himalayan ranges, *Earth and Planetary Science Letters*, *220*, 379–389.

Hsu, L., and R. Bürgmann (2006), Surface creep along the Longitudinal Valley fault, Taiwan from InSAR measurements, *Geophysical Research Letters*, *33, L06312*, doi: 10.1029/2005GL024624.

Hsu, Y.-J., M. Simons, J.-P. Avouac, J. Galetzka, K. Sieh, C. M, D. Natawidjaja, L. Prawirodirdjo, and Y. Bock (2006), Frictional afterslip following the Mw 8.7, 2005 Nias-Simeulue earthquake, Sumatra, *Science*, *312*, 1921–1926.

Hsu, Y.-J., S.-B. Yu, and H.-Y. Chen (2009a), Coseismic and postseismic deformation associated with the 2003 Chengkung, Taiwan, earthquake, *Geophysical Journal International*, *176(2)*, 420–430.

Hsu, Y.-J., J.-P. Avouac, S.-B. Yu, C.-H. Chang, Y.-M. Wu, and J. Woessner (2009b), Spatio-temporal Slip, and Stress Level on the Faults within the Western Foothills of Taiwan: Implications for Fault Frictional Properties, *Pure and Applied Geophysics*, *166*, 1853–1884.

Ide, S. (2010), Striations, duration, migration and tidal response in deep tremor, *Nature*, *466*, 356–359.

Ide, S., D. R. Shelly, and G. C. Beroza (2007), Mechanism of deep low frequency earthquakes: Further evidence that deep non-volcanic tremor is generated by shear slip on the plate interface, *Geophysical Research Letters*, *34, L03308*, doi:10.1029/2006GL028890.

International Seismological Centre (2010), *On-line Bulletin*, Int. Seis. Cent., Thatcham, United Kingdom, http://www.isc.ac.uk.

Jackson, M., and R. Bilham (1994), Constraints on Himalayan deformation inferred from vertical velocity fields in Nepal and Tibet, *Journal of Geophysical Research*, *99*, 13,897–13,912.

Jouanne, F., J. Mugnier, M. Pandey, J. Gamond, P. Le Fort, L. Serrurier, C. Vigny, and J.-P. Avouac (1999), Oblique convergence in the Himalayas of Western Nepal Deduced from Preliminary Results of GPS Measurements, *Geophysical Research Letters*, *26(13)*, 1933–1936.

Jouanne, F., J. Mugnier, J. Gamond, P. Le Fort, M. Pandey, L. Bollinger, M. Flouzat, and J. Avouac (2004), Current shortening across the Himalayas of Nepal, *Geophysical Journal International*, *157*, 1–14.

Jouanne, F., A. Awan, A. Madji, A. Pêcher, M. Latif, A. Kausar, J. L. Mugnier, I. Khan, and N. A. Khan (2011), Postseismic deformation in Pakistan after the 8 October 2005 earthquake: Evidence of afterslip along a flat north of the Balakot-Bagh thrust, *Journal of Geophysical Research*, *116, B07401*, doi:10.1029/2010JB007903.

Kaneko, Y., and N. Lapusta (2008), Variability of earthquake nucleation in continuum models of rate-and-state faults and implications for aftershock rates, *Journal of Geophysical Research*, *113, B12312*, doi:10.1029/2007JB005154.

Kaneko, Y., J.-P. Avouac, and L. Nadia (2010), Towards inferring earthquake patterns from geodetic observations of interseismic coupling, *Nature Geoscience*, *3*, 363–369, doi: 10.1038/ngeo843.

Knott, C. (1897), On Lunar Periodicities in Earthquake Frequency, *Proceedings of the Royal Society of London*, *60*, 457–466.

Konca, O., J.-P. Avouac, A. Sladen, A. J. Meltzner, K. Sieh, P. Fang, Z. Li, J. Galetzka, J. Genrich, M. Chlieh, D. H. Natawidjaja, Y. Bock, E. J. Fielding, C. Ji, and D. V. Helmberger (2008), Partial rupture of a locked patch of the Sumatra megathrust during the 2007 earthquake sequence, *Nature*, *456*, 631–635.

Kumar, S., S. Wesnousky, T. Rockwell, R. Briggs, V. Thakur, and R. Jayangondaperumal (2006), Paleoseismic evidence of great surface rupture earthquakes along the Indian Himalaya, *Journal of Geophysical Research*, *111, B03304*, doi:10.1029/2004JB003309.

Kumar, S., S. Wesnousky, R. Jayangondaperumal, T. Nakata, Y. Kumahara, and V. Singh (2010), Paleoseismological evidence of surface faulting along the northeastern Himalayan front, India: Timing, size, and spatial extent of great earthquakes, *Journal of Geophysical Research*, *115, B12422*, doi:10.1029/2009JB006789.

Langbein, J., and H. Johnson (1997), Correlated errors in geodetic time series: Implications for time-dependent deformation, *Journal of Geophysical Research*, *102*, 591–603.

Lapusta, N., and Y. Liu (2009), Three-dimensional boundary integral modeling of spontaneous earthquake sequences and aseismic slip, *Journal of Geophysical Research*, *114, B09303*, doi:doi:10.1029/2008JB005934.

Lapusta, N., and J. R. Rice (2003), Nucleation and early seismic propagation of small and large events in a crustal earthquake model, *Journal of Geophysical Research*, *108, NO. B4*, 2205, doi:10.1029/2001JB000793.

Lapusta, N., J. R. Rice, G. BenZion, and G. Zheng (2000), Elastodynamic analysis for slow tectonic loading with spontaneous rupture episodes on faults with rate- and state-dependent friction, *Journal of Geophysical Research*, *105*, 23,765–23,789, doi: 10.1029/2000JB900250.

Larson, K. M., R. Bürgmann, R. Bilham, and J. T. Freymueller (1999), Kinematics of the India-Eurasia collision zone from GPS measurements, *Journal of Geophysical Research*, *104*, 1077–1093.

Lavé, J., and J.-P. Avouac (2000), Active folding of fluvial terraces across the Siwaliks Hills, Himalayas of central Nepal, *Journal of Geophysical Research - Solid Earth*, *105*, 5735–5770.

Lavé, J., D. Yule, S. Sapkota, K. Basant, C. Madden, M. Attal, and R. Pandey (2005), Evidence for a great medieval earthquake (approximate to 1100 AD) in the Central Himalayas, Nepal, *Science*, *307*, 1302–1305.

Lockner, D., and N. Beeler (1999), Premonitory slip and tidal triggering of earthquakes, *Journal of Geophysical Research - Solid Earth*, *104*(B9), 20,133–20,151.

Loveless, J., and B. Meade (2010), Geodetic imaging of plate motions, slip rates, and partitioning of deformation in Japan, *Journal of Geophysical Research-Solid Earth*, *115, B02410*, doi:10.1029/2008JB006248.

Lowry, A. R. (2006), Resonant slow fault slip in subduction zones forced by climatic load stress, *Nature*, *442*, 802–805.

Marone, C. (1998), Laboratory-derived friction laws and their application to seismic faulting, *Ann. Revs. Earth & Plan. Sci.*, *26*, 643–696.

Meade, B. J. (2010), The signature of an unbalanced earthquake cycle in Himalayan topography?, *Geology*, *38 (11)*, 987–990.

Melbourne, T. I., F. H. Webb, J. M. Stock, and C. Reigber (2002), Rapid postseismic transients in subduction zones from continuous GPS, *Journal of Geophysical Research-Solid Earth*, *107 (B10)*, 2241, doi:10.1029/2001JB000555.

Métois, M., A. Socquet, and C. Vigny (2012), Interseismic coupling, segmentation and mechanical behavior of the central Chile subduction zone, *Journal of Geophysical Research*, *117, B03406*, doi:10.1029/2011JB008736.

Miyazaki, S., P. Segall, J. Fukuda, and T. Kato (2004), Space time distribution of afterslip following the 2003 Tokachi-oki earthquake: Implications for variations in fault zone frictional properties, *Geophysical Research Letters*, *31, L06623*, doi:10.1029/2003GL019410.

Miyazawa, M., and E. E. Brodsky (2008), Deep low-frequency tremor that correlates with passing surface waves, *Journal of Geophysical Research*, *113, B01307*, doi: 10.1029/2006JB004890.

Molnar, P. (1979), Earthquake recurrence intervals and plate tectonics, *Bulletin of the Seismological Society of America*, *69*, 115–133.

Molnar, P. (1987), The distribution of intensity associated with the 1905 Kangra earthquake and bounds on the extent of the rupture zone., *J. Geol. Soc. of India*, *29*, 211–229.

Moore, D., D. Lockner, S. Ma, R. Summers, and J. Byerlee (1997), Strengths of serpentinite gouges at elevated temperatures, *Journal of Geophysical Research*, *102(B7)*, 14,787–14,801.

Moreno, M., M. Rosenau, and O. Onno (2010), 2010 Maule earthquake slip correlates with pre-seismic locking of Andean subduction zone, *Nature*, *467*, 198–202, doi: 10.1038/nature09349.

Mugnier, J., P. Huyghe, P. Leturmy, and F. Jouanne (2003), Episodicity and Rates of Thrust-sheet Motion in the Himalayas (Western Nepal), *AAPG Mem*, *82*, 1–24.

Nakata, R., N. Suda, and H. Tsuruoka (2008), Non-volcanic tremor resulting from the combined effect of Earth tides and slow slip events, *Nature*, *1*, 676–678, doi: doi:10.1038/ngeo288.

Nakata, T. (1989), Active faults of the Himalayas of India and Nepal, *Geol. Soc. Am., spec.*, *232*, 243–264.

Noda, H., and N. Lapusta (2010), Three-dimensional earthquake sequence simulations with evolving temperature and pore pressure due to shear heating: Effect of heterogeneous hydraulic diffusivity, *Journal of Geophysical Research*, *115, B12314*, doi: 10.1029/2010JB007780.

Nur, A., and J. R. Booker (1972), Aftershocks Caused by Pore Fluid Flow?, *Science*, *175*, 885–887.

Okada, Y. (1985), Surface deformation due to shear and tensile faults in a half-space, *Bulletin of the Seismological Society of America*, *75(4)*, 1135–1154.

Ozawa, S., T. Nishimura, H. Suito, T. Kobayashi, M. Tobita, and T. Imakiire (2011), Coseismic and postseismic slip of the 2011 magnitude-9 Tohoku-Oki earthquake, *Nature*, *475*, 373–376, doi:10.1038/nature10227.

Palmer, A. C., and J. R. Rice (1973), The growth of slip surfaces in the progressive failure of overconsolidated clay slopes, *Proceedings of the Royal Society of London*, *332*, 527–548.

Pandey, M., R. Tankudar, J.-P. Avouac, J. Lavé, and J.-P. Massot (1995), Interseismic strain accumulation on the Himalayan crustal ramp (Nepal), *Geophysical Research Letters*, *22*, 751–754.

Pandey, M., R. Tankudar, J.-P. Avouac, J. Vergne, and T. Héritier (1999), Seismotectonics of the Nepal Himalaya from a local seismic network, *Journal of Asian Earth Sciences*, *17*, 703–712.

Parsons, T. (2004), A hypothesis for delayed dynamic earthquake triggering, *Geophysical Research Letters*, *32, L04302*, doi:10.1029/2004GL021811.

Perfettini, H., and J.-P. Avouac (2004), Postseismic relaxation driven by brittle creep: A possible mechanism to reconcile geodetic measurements and the decay rate of aftershocks, application to the Chi-Chi earthquake, Taiwan, *Journal of Geophysical Research*, *109 (B2), B02304*, doi:10.1029/2003JB002488.

Perfettini, H., and J.-P. Avouac (2007), Modeling afterslip and aftershocks following the 1992 Landers earthquake, *Journal of Geophysical Research*, *112, B07409*, doi: 10.1029/2006JB004399.

Perfettini, H., and J. Schmittbuhl (2001), Periodic Loading On A Creeping Fault: Implications For Tides, *Geophysical Research Letters*, *28(3)*, 435–438, doi: 10.1029/2000GL011686.

Perfettini, H., J. Schmittbuhl, J. R. Rice, and M. Cocco (2001), Frictional response induced by time-dependent fluctuations of the normal loading, *Journal of Geophysical Research*, *106(B7)*, 13,455–13,472, doi:10.1029/2000JB900366.

Perfettini, H., J.-P. Avouac, H. Tavera, A. Kositsky, J.-M. Nocquet, F. Bondoux, M. Chlieh, A. Sladen, L. Audin, D. L. Farber, and P. Soler (2010), Seismic and aseismic slip on the Central Peru megathrust, *Nature*, *465 (7294)*, 78–81, doi:10.1038/nature09062.

Ponraj, M., S. Miura, C. D. Reddy, S. Amirtharaj, and S. H. Mahajana (2011), Slip distribution beneath the Central and Western Himalaya inferred from GPS observations, *Geophysical Journal International*, *185*, 724–736.

Press, W., S. Teukolsky, W. Vetterling, and B. Flannery (1992), *Numerical Recipes in C: The Art of Scientific Computing*, New York: Cambridge University Press.

Proksch, A., and H. R. Baidya (1995), *Images of a Century, the Changing Townscapes of the Kathmandu Valley*, Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ), Urban Development Through Local Efforts Project (ULDE).

Rayleigh, L. (1880), On the result of a large number of vibrations of the same pitch and arbitrary phase, *Phil. Mag.*, *10*, 73–78.

Reasenberg, P. (1985), Second-Order Moment of Central California Seismicity, 1969-1982, *Journal of Geophysical Research*, *90*, 5479–5495.

Rice, J. R., and A. Ruina (1983), Stability of steady frictional slipping, *Journal of Applied Mechanics*, *50*, 343–349.

Rogers, G., and H. Dragert (2003), Episodic tremor and slip on the Cascadia subduction zone: The chatter of silent slip, *Science*, *300(5627)*, 1942–1943.

Rubin, A. M., and J.-P. Ampuero (2005), Earthquake nucleation on (aging) rate and state faults, *Journal of Geophysical Research*, *110, B11312*, doi:10.1029/2005JB003686.

Rubinstein, J. L., M. L. Rocca, J. E. Vidale, K. C. Creager, and A. G. Wech (2008), Tidal Modulation of Nonvolcanic Tremor, *Science*, *319(5860)*, 186–189.

Ruina, A. (1983), Slip instability and state variable friction laws, *Journal of Geophysical Research*, *88*, 10,359–10,370, doi:10.1029/JB088iB12p10359.

Rydelek, P. A., and L. Hass (1994), On Estimating the Amount of Blasts in Seismic Catalogs with Schuster's Method, *Bulletin of the Seismological Society of America*, *84(4)*, 1256–1259.

Sapkota, S. N., L. Bollinger, Y. Klinger, P. Tapponnier, Y. Gaudemer, and D. R. Tiwari (2013), Primary surface ruptures of the great Himalayan earthquakes in 1934 and 1255, *Nature Geoscience*, *6*, 71–76.

Savage, H. M., and C. Marone (2007), Effects of shear velocity oscillations on stick-slip behavior in laboratory experiments, *Journal of Geophysical Research*, *112, B02301*, doi:10.1029/2005JB004238.

Savage, H. M., and C. Marone (2008), Potential for earthquake triggering from transient deformations, *Journal of Geophysical Research*, *113, B05302*, doi:10.1029/2007JB005277.

Savage, J. (1983), A dislocation model of strain accumulation and release at a subduction zone, *Journal of Geophysical Research*, *88*, 4984–4996, doi:10.1029/JB088iB06p04984.

Scargle, J. D. (1982), Studies in astronomical time series analysis. II. Statistical aspects of spectral analysis of unevenly spaced data, *The Astrophysical Journal*, *263*, 835–853.

Schuster, A. (1897), On Lunar and Solar Periodicities of Earthquakes, *Proceedings of the Royal Society of London*, *61*, 455–465.

Schwartz, S., and J. Rokosky (2007), Slow slip events and seismic tremor at circum-Pacific subduction zones, *Reviews of Geophysics*, *45 (3)*, doi:10.1029/2006RG000208.

Seeber, L., and V. Gornitz (1983), River profiles along the Himalayan arc as indicators of active tectonics, *Tectonophysics*, *92 (4)*, 335–337, 341–367, doi:10.1016/0040-1951(83)90201-9.

Segall, P. (2010), *Earthquake and Volcano Deformation*, Princeton University Press.

Shelly, D. R., Z. Peng, D. P. Hill, and C. Aiken (2011), Triggered creep as a possible mechanism for delayed dynamic triggering of tremor and earthquakes, *Nature geoscience*, *4*, 384–388.

Shibazaki, B., and Y. Iio (2003), On the physical mechanism of silent slip events along the deeper part of the seismogenic zone, *Geophysical Research Letters*, *30(9)*, 1489, doi:10.1029/2003GL017047.

Shibazaki, B., and T. Shimamoto (2007), Modelling of short-interval silent slip events in deeper subduction interfaces considering the frictional properties at the unstable-stable transition regime, *Geophysical Journal International*, *171*, 191–205.

Shimamoto, T. (1986), Transition between frictional slip and ductile flow for halite shear zones at room temperature, *Science*, *231*, 711–714.

Socquet, A. (2003), Accomodation du mouvement relatif entre l'Inde et le Sonde: depuis la Faille de Sagaing jusqu'à la Syntaxe Est Himalayenne, Ph.D. thesis, Université Paris XI.

Styron, R., M. Taylor, and M. Murphy (2011), Oblique convergence, arc-parallel extension, and the role of strike-slip faulting in the High Himalaya, *Geosphere*, *7*, 582–596.

Suwa, Y., S. Miura, A. Hasegawa, T. Sato, and K. Tachibana (2006), Interplate coupling beneath NE Japan inferred from three-dimesional displacement field, *Journal of Geophysical Research*, *111*, doi:10.1029/2004JB003203.

Tanaka, S., M. Ohtake, and H. Sato (2002a), Evidence for tidal triggering of earthquakes as revealed from statistical analysis of global data, *Journal of Geophysical Research*, *107*, 2211, doi:10.1029/2001JB001577.

Tanaka, S., M. Ohtake, and H. Sato (2002b), Spatio-temporal variation of the tidal triggering effect on earthquake occurrence associated with the 1982 South Tonga earthquake of M_W 7.5, *Geophysical Research Letters*, *29(16)*, doi:10.1029/2002GL015386.

Tanaka, S., H. Sato, S. Matsumura, and M. Ohtake (2006), Tidal triggering of earthquakes in the subducting Philippine Sea plate beneath the locked zone of the plate interface in the Tokai region, Japan, *Tectonophysics*, *417*, 69–80.

Thomas, A. M., R. M. Nadeau, and R. Bürgmann (2009), Tremor-tide correlations and near-lithostatic pore pressure on the deep San Andreas fault, *Nature*, *462*, 1048–1051, doi:10.1038/nature08654.

Thomas, A. M., R. Bürgmann, D. R. Shelly, N. M. Beeler, and M. L. Rudolph (2012), Tidal triggering of low frequency earthquakes near Parkfield, CA: implications for fault mechanics within the brittle-ductile transition, *Journal of Geophysical Research - Solid Earth*, *117, B05301*, doi:10.1029/2011JB009036.

Toda, S., R. S. Stein, P. A. Reasenberg, J. H. Dieterich, and A. Yoshida (1998), Stresstransferred by the 1995 $M_w = 6.9$ Kobe, Japan, shock: Effect on aftershocks and future earthquake probabilitie, *Journal of Geophysical Research*, *103, NO. B10*, 24,543–24,565.

Toda, S., R. S. Stein, and T. Sagiya (2002), Evidence from the AD 2000 Izu islands earthquake swarm that stressing rate governs seismicity, *Nature Geoscience*, *419*, 58–61.

Toda, S., R. S. Stein, G. C. Beroza, and D. Marsan (2012), Aftershocks halted by static stress shadows, *Nature Geoscience*, *5*, 410–413, doi:doi:10.1038/ngeo1465.

Utsu, T., Y. Ogata, and R. S. Matsu'ura (2005), The centenary of the Omori formula for a decay law of aftershock activity, *Journal of Physics of the Earth*, *43*, 1–33.

Vergne, J., R. Cattin, and J.-P. Avouac (2001), On the use of dislocations to model interseismic strain and stress build-up at intracontinental thrust faults, *Geophysical Journal International*, *147*, 155–162.

Waldhauser, F., and W. L. Ellsworth (2000), A Double-Difference Earthquake Location Algorithm: Method and Application to the Northern Hayward Fault, California, *Bulletin of the Seismological Society of America*, *90*, 1353–1368.

Wallace, L., J. Beavan, R. McCaffrey, and D. Darby (2004), Subduction zone coupling and tectonic block rotations in the North Island, New Zealand, *Journal of Geophysical Research-Solid Earth*, *109 (B12)*, doi:10.1029/2004JB003241.

Williams, S. D., Y. Bock, P. Fang, P. Jamason, R. M. Nikolaidis, L. Prawirodirdjo, M. Miller, and D. J. Johnson (2004), Error analysis of continuous GPS position time series, *Journal of Geophysical Research*, *109, B03412*, doi:10.1029/2003JB002741.

Williams, S. D. P. (2003a), The effect of coloured noise on the uncertainties of rates estimated from geodetic time series, *Journal of Geodesy*, *76*, 483–494, doi:10.1007/s00190-002-0283-4.

Zhang, J., Y. Bock, H. Johnson, P. Fang, S. Williams, J. Genrich, S. Wdowinski, and J. Behr (1997), Southern California Permanent GPS Geodetic Array: Error analysis of daily position estimates and site velocities, *Journal of Geophysical Research*, *102(B8)*, 18,035–18,055, doi:10.1029/97JB01380.

Zhao, W., K. Nelson, and P. I. team (1993), Deep seismic reflection evidence for continental underthrusting beneath southern Tibet, *Nature*, *366*, 555–559, doi:10.1038/366557a0.