Abstract

One way to probe earthquake nucleation processes and the relation between stress buildup and seismicity is to analyze the sensitivity of seismicity to stress perturbations. Here, we report evidence for seasonal strain and stress (~2–4 kPa) variations in the Nepal Himalaya, induced by water storage variations which correlate with seasonal variations of seismicity. The seismicity rate is twice as high in the winter as in the summer, and correlates with stress rate variations. We infer ~10–20 kPa/yr interseismic stress buildup within the seismicity cluster along the high Himalaya front. Given that Earth tides exert little influence on Himalayan seismicity, the correlated seasonal variation of stress and seismicity indicates that the duration of earthquake nucleation in the Himalaya is of the order of days to month, placing constraints on faults friction laws. The unusual sensitivity of seismicity to small stress changes in the Himalaya might be due to high pore pressure at seismogenic depth.

© 2007 Elsevier B.V. All rights reserved.

Keywords: seismology; geodesy; tectonics; hydrology; Himalaya

1. Introduction

Background seismicity in the Himalaya is known to be driven by the slow, presumably steady-state, process of interseismic stress buildup in the period separating very large (M>8) earthquakes (Cattin and Avouac, 2000; Bollinger et al., 2004). However, strong seasonal fluctuations are observed, with winter seismicity twice as high as summer seismicity, on average (Bollinger et al., 2007). We analyze continuous geodetic data in order to investigate the cause of these variations, and derive implications for earthquake nucleation and fault rheology. Indeed, the earthquake nucleation process determines how seismicity responds to temporal or spatial variations in strain, whether these are related to coseismic deformation, postseismic relaxation, interseismic stress buildup, or periodic perturbations such as Earth tides (Dieterich et al., 2000; Dieterich, 1994, 1987; Lockner and Beeler,
2. Background seismicity, secular strain and crustal structure

Local seismic monitoring in the Nepal Himalaya has revealed very intense background seismicity along the front of the high range (Panney et al., 1995, 1999) (Fig. 1). Most of the activity comes from thrust events induced by north–south compression related to the ongoing convergence across the Himalaya. The secular velocities derived from geodetic measurements across the Nepal Himalaya show that the Main Himalayan Thrust (MHT) is locked from beneath the high range to the piedmont, where it surfaces (Fig. 1), and that it roots to the north into a subhorizontal shear zone that is creeping at about 2 cm/yr (Fig. 2a) (Bilham et al., 1997; Bettinelli et al., 2006). Most of the seismicity clusters near the updip edge of the creeping zone (Fig. 2b), in a region where, according to the modeling of the geodetic data, Coulomb stress builds up at a rate higher than about 6 kPa/yr (Cattin and Avouac, 2000; Bollinger et al., 2004). This midcrustal seismicity was also observed to coincide with a...
3. Seasonality of strain and seismicity

The geodetic time series (Figs. 2b, 3b and c) and seismicity (Fig. 3a) both show strong seasonal variations. Interpretation of the variations requires some care, as they could result from a variety of meteorological...

---

**Fig. 2.** Secular interseismic strain and seasonal variation across Nepal Himalaya. a Secular velocities relative to India determined from campaign measurements (black dots) and from the analysis of the time series recorded at the continuous geodetic stations (Bettinelli et al., 2006) (red dots, uncertainties are smaller than the dots size). Continuous line shows the predicted horizontal velocity for a slip rate of 16 mm/yr, a shallow dipping dislocation as indicated in cross section and derived from least-squares adjustments of all GPS and leveling data from central Nepal (Bettinelli et al., 2006). b GPS time series de-trended by removing the secular motion. The seasonal variations increase in amplitude from south to north. Details on the processing of the GPS data are given in Bettinelli et al., (2006). Red shaded areas show seasonal geodetic displacements computed from the variation of surface load due to seasonal variation of water storage (2 sigma uncertainty) assuming an elastic half space with a elastic shear modulus of 40 GPa and a Poisson coefficient of 0.25. c simplified geological cross section and geophysical constraints on the crustal structure across central Nepal. See Fig. 1 for location of section. The conductivity section was obtained from a magnetotelluric experiment carried out along section AB across central Nepal (Lemmonier et al, 1999). Also reported are the INDEPTH seismic sections about 500 km east of section AB (Brown et al, 1996; Nelson et al, 1996; Zhao et al., 1993). All thrust faults are inferred to root at depth into a subhorizontal ductile shear zone that coincides with a prominent midcrustal reflector. White circles show seismicity corresponding to events with well-constrained hypocentral depths (Cattin and Avouac, 2000).
artifacts: the GPS time series might be biased by mis-correction of the tropospheric delays and seasonal variations in the seismic noise level could affect the detection threshold of the seismic network.

Analysis of a catalogue of ~10,000 earthquakes (Bollinger et al., 2007) shows that the seismicity rate in the winter is twice as high as in the summer at all magnitudes above the detection threshold (estimated to be MI = 2.0 (Pandey et al., 1999)), which rules out the possibility that the pattern arises solely from seasonal variations in seismic noise. This analysis also shows an extremely low probability—only $10^{-15}$ when computed with the Schuster test as applied to the declustered catalogue (Schuster, 1897; Heaton, 1975)—that this seasonality is due to chance. We have analyzed potential tropospheric bias in the GPS measurements. The details of this analysis is presented elsewhere (Bettinelli, 2007) and only summarized briefly here. We have first compared the troposphere delays estimated from the GPS data with those predicted from the only in situ meteorological data, and verify that, although these latter are not representative of the whole atmospheric
column, there is no significant discrepancy in the overall shape and amplitude (Bettinelli, 2007) (Fig. S2). The vertical positions, which are known to be particularly sensitive to tropospheric effects, do however show some correlation with the seasonal variations of tropospheric delays suggesting possible atmospheric bias. By contrast, the seasonal variations of horizontal positions are in phase neither with the tropospheric delays, the phase shift being of about 4 months, nor with the variations of vertical positions (Bettinelli, 2007). Temporal variations in the horizontal gradient of the tropospheric delay cannot be either the cause of the observed seasonal variations of horizontal positions since they also vary in phase with the tropospheric delays (Bettinelli, 2007). The horizontal strain variations derived from our GPS measurements must thus represent real ground deformation. Variations of the vertical component are less reliable. They probably reflect a combination of real ground displacement and tropospheric artifacts.

The amplitude of the horizontal seasonal geodetic displacements increases northwards, so that the shortening rate across the range is lower in the summer than in the winter. This points to a local source capable of producing horizontal contraction in the winter and horizontal extension in the summer, superimposed on the secular shortening rate across the Himalaya.

4. Potential causes of seasonal strain in the Himalaya

Seasonal variations not related to tropospheric mis-corrections are commonly observed in geodetic time series and may have a variety of causes (Blewitt and Lavallee, 2002; Bawden et al., 2001; Dong et al., 2002; Ben-Zion and Leary, 1986) but only few can produce the kind of seasonal strain variation reported in this study, and explain the correlated variations of geodetic strain and seismicity. One possibility is that the creep rate on the MHT, which drives interseismic shortening, varies with time due to a resonance with atmospheric seasonal forcing (Perfettini and Schmittbuhl, 2001). Such a mechanism has been invoked to explain the periodicity of slow events (Lowry, 2006; Shen et al., 2005). Another possibility is that the creep rate on the MHT is constant, but a seasonal source of loading produces stress variations capable of modulating the seismicity and geodetic strain. Thermoelastic strain has been shown to be a dominant source of seasonal strain in California (Ben-Zion and Leary, 1986; Prawiradiljo et al., 2006) and could be invoked in the Nepal Himalaya as well. Atmospheric pressure, hydrology and snow cover can also all produce seasonal variations in surface loading, which could in turn cause seasonal variations in geodetic strain and seismicity. In Japan, for example, snow load variations produce observable seasonal geodetic displacements, and seasonal seismicity (Heki, 2003). In California, atmospheric pressure has been proposed to explain seasonal variations in seismicity following the Landers earthquake (Gao et al., 2000). In the Himalaya, snow load and atmospheric pressure are both at a maximum in the winter (Putkonen, 2004). These effects should, contrary to what is observed, inhibit thrust earthquakes in the winter (Bollinger et al., 2007).

Variation of surface loading associated with the hydrological cycle is another possible mechanism that has been shown to be capable of producing measurable geodetic strain at a regional scale (Blewitt et al., 2001; van Dam et al., 2001). We have explored more particularly this potential effect using satellite altimetry measurements from TOPEX-Poseidon and geoid data from the GRACE mission (Fig. 3d) to estimate the temporal variations of land water storage (Frappart et al., 2006). We selected TOPEX-Poseidon altimetric measurements over major rivers within the Ganges basin (Fig. 1). These measurements provide an estimate of the free surface water level to within a few tens of cm. They show that during the monsoon season the free water level over the whole Gangetic basin rises by about 4 m, starting in mid-May, reaches a maximum in early September, followed by a steady decrease until the next monsoon. Strong, in-phase, seasonal variations in seismicity, geodetic displacements and water levels in the Ganges basin suggest a striking causal relationship (Fig. 3).

The GRACE satellite measurements of the geopotential complemented the TOPEX-Poseidon observations, which ended in 2002. In addition the GRACE measurements provide an estimate of the mass surface loading due to free surface water, groundwater and soil moisture. The GRACE geoid data set consists of monthly estimates of time series of spherical harmonics up to degree and order 50, corresponding to a spatial resolution of 400 km at the surface of the Earth (Tapley et al., 2004; Tapley et al., 2004; Schmidt et al., 2006). In the process, atmospheric and oceanic tide effects are removed from the geoid harmonic coefficients. The coefficients of the contribution of continental water mass storage, including surface waters, soil moisture and groundwater, are extracted from the 10-day sampled GRACE solutions (Lemoine et al., 2007) using an iterative inverse method (Ramillien et al., 2006, 2005). Strong seasonal variations of surface load are revealed. Variation of water storage in the Ganges basin dominates the signal with a peak-to-peak amplitude of about 50 cm (Figs. 3d and 4). The radar altimetry and GRACE-derived land water time series do not overlap, but a similar phase is found in the seasonal cycle and the
amplitude of the equivalent water thickness derived from GRACE is about ten times smaller than the variations of rivers surface elevation (Fig. 3). The free surface within the major rivers and the groundwater level vary jointly over the seasonal cycle, probably due to the relatively high permeability of the clastic sediments of the Ganges basin (Fig. 3). The surface load variations over the whole period can thus be estimated by combining the radar altimetry and the GRACE-derived land water time.

5. Influence of water storage variations on strain and stress in the Himalaya

We propose that the seasonal geodetic displacements reflect a lithospheric response to the seasonal variation of hydrological surface loads (Fig. 5). This mechanism is qualitatively consistent with our geodetic observations since the surface load rises during the summer monsoon is expected to induce extension across the Himalaya, while compression should follow in the winter when the surface load decreases. For a more quantitative assessment, we modeled this effect using two different approaches. The first is based on the analytical solution of Boussinesq (1885) for deformation of an elastic half-space submitted to a surface point load (Fig. 6), an approach used in previous similar study (Grape nthin et al., 2006). This model ignores the effects of topography and of viscous deformation. The second approach uses a finite element model in which a purely elastic plate is assumed to overly a nonviscous fluid (Figs. S1 and S2). This modeling takes topography into account. In both cases, a load mimicking the effects of water level variation is applied at the surface. The spatial distribution of surface load was estimated from the

![Estimated seasonal amplitudes from Grace Land Water Solutions](image)

Fig. 4. Map of the seasonal peak-to-peak amplitudes of total water storage in the Southeast region, which were estimated from the GRACE Land Waters solutions up to harmonic degree 50 by linear leastsquare inversion (units: meters of equivalent-water thickness). Note the important annual variations of water mass over the Mekong and the Ganges basins that reach ±2.4 m of water height ([http://www.legos.obs-mip.fr/en/equipes/gohs/resultats/c1_grace1](http://www.legos.obs-mip.fr/en/equipes/gohs/resultats/c1_grace1)). The dots show the locations of the GPS stations analyzed in this study.
GRACE data and the temporal variations were estimated from combining the TOPEX-Poseidon and GRACE data (Fig. 3). The predicted seasonal displacements of both modeling approaches match the observations remarkably well (Figs. 3, 6 and S2). Although other factors might contribute to some degree, variations of water storage in the Ganges basin turns out to be the primary cause of the observed seasonality of geodetic displacements.

6. Relation between seasonal variations of stress and seismicity

This mechanism is also qualitatively consistent with the higher seismicity rate in the winter, suggesting that the gradual decrease of surface load in the Ganges basin induces a gradual increase of compression in the Himalaya, which adds to the effect of interseismic stress accumulation at the front of the creeping portion of the MHT (Fig. 6). To quantify the effect of stress variations on seismicity, we use the Coulomb stress, defined as

\[ S = \tau - \mu' \cdot \sigma, \]

where \( \tau \) is the shear stress, \( \sigma \) is the normal stress and \( \mu' \) is an effective friction coefficient, here set to 0.3. The Coulomb stress is commonly used to evaluate whether stress variations should promote or inhibit seismicity (King et al., 1994). We find that seasonal variations in the water level of the Ganges basin induce fluctuations in the Coulomb stress computed for the zone of seismicity along the front of the high range, with a peak-to-peak amplitude of 2–4k Pa (Fig. 7). In fact, the seismicity rate is found to correlate
best with stress rate (Figs. 8 and 9), indicating that triggering of earthquakes depends on the stressing rate rather than on the absolute stress level, as has been observed in a number of previous case studies (Toda et al., 2002; Perfettini and Avouac, 2004; Hsu et al., 2006).

Fig. 6. Observed and predicted NS, EW and vertical displacements at stations SIMR, DAMA, GUMB, NAGA, EVEB and LHAS. Elastic half-space. The theoretical displacements were computed by convolving the distribution of surface load with the analytical solution for the deformation of an elastic half space submitted to a surface point load assuming a Young modulus of 40 GPa. [The equations originally from Boussinesq (1885) are given in standard text books on linear elasticity and can be found for example at http://www.engin.brown.edu/courses/en224/halfspace/halfspace.html]. The effect of the topography is neglected in this computation, as well as that of viscous deformation. The blue shaded areas show the seasonal geodetic displacements computed from the surface load exerted by seasonally varying water storage taking into account the spread in the GRACE and TOPEX-Poseidon measurements (shown in Fig. 3d).
Assuming that this linear correlation has some physical significance the sensitivity of seismicity rate to seasonal stress fluctuations can be used to estimate the secular interseismic stress rate independently of the modeling of interseismic strain. To do so, we have solved for the secular stress rate which yields the best linear correlation between the stress rate and the seismicity rate and obtained a value between 10 and 20 kPa/yr (Fig. 9). This rate represents the average stress rate within the seismicity zone along the front of the high Himalaya where stress builds up in the interseismic period around the downdip end of the locked portion of the Main Himalayan Thrust fault (Cattin and Avouac, 2000; Bollinger et al., 2004). It is likely, however, that the stress rate within this zone is highly variable depending on the geometry and complexity of the transition between the locked and creeping fault zones. The agreement between the estimate of the stress rate derived from the seasonality of seismicity and that estimated from the modeling of secular interseismic strain (>6 kPa/yr (Cattin and Avouac, 2000; Bollinger et al., 2004)) is remarkable in light of the uncertainties and approximations made in these two independent approaches.

**7. Implications for earthquake nucleation**

The sensitivity of Himalayan seismicity to seasonal stress and strain variations has important implications for earthquake nucleation.

We first consider a simple Coulomb failure model in which rupture occurs instantaneously when S first reaches some critical value $S_c$ and then drops to some dynamic value $S_d$. In this case, if the initial Coulomb stress on various faults is random distributed between $S_d$ and $S_c$, and if the fault population is subjected to a constant stress...
rate $\dot{S}_0$, then the seismicity rate $r_0$ is proportional to $\dot{S}_0$. Let’s now consider that the system is perturbed by Coulomb stress oscillations with period $T$ and amplitude $S_m$ small enough that the stress rate is always positive (i.e. $\dot{S}_0 \geq 2\pi \frac{S_m}{T}$), as is the case in this study. The seismicity rate is then expected to vary in proportion to the stress rate $\dot{S}$, with periodic variations of amplitude $R_m$ given by:

$$R_m = 2\pi \frac{S_m}{T \cdot \dot{S}_0} \cdot r_0$$  \hspace{1cm} (1)

This model makes it explicit that the secular stress rate, $\dot{S}_0$, can be estimated from observations of periodic variations in seismicity, and, as argued in the previous section, accounts reasonably well for the response of Himalayan seismicity to annual stress variations. Eq. (1) also suggests that the sensitivity of seismicity to periodic stress variations should increase as the period decreases, such that seismicity should be extremely sensitive to Earth tides. Indeed, Earth tides induce variations in stress on the order of 3–4 kPa, an amplitude similar to those related to the seasonal variation of surface water storage, but at a much shorter dominant period of 12 h. As already reported (Cochran et al., 2004), no clear correlation between seismicity and Earth tides is actually observed in the Himalaya however. This suggests that the response of seismicity to periodic stress perturbations is frequency dependent. This is most probably due to the fact that, at the time scale of Earth tides, the duration of earthquake nucleation is not negligible and a simple Coulomb failure model therefore does not apply (Dieterich, 1987; Lockner and Beeler, 1999; Beeler and Lockner, 2003). The duration of nucleation of earthquakes in the Himalaya is thus probably short compared to the seasonal time scale but long compared to the semi-diurnal and 14 days dominant periods of Earth tides.

Some quantitative constraints on the constitutive laws of faults might be derived from this analysis. Within the theoretical framework of rate-and-state friction (Dieterich, 1994, 1987; Beeler and Lockner, 2003; Perfettini et al., 2003), seismic rupture is preceded by a nucleation phase of self-accelerating slip with a duration characterized by

$$t_a = \frac{a\sigma}{\dot{S}},$$  \hspace{1cm} (2)

where $\sigma$ is the normal stress and $a$ is a constitutive parameter reflecting the dependence of the friction coefficient, $\mu = \frac{2}{3}$, on slip velocity, $V$, according to

$$a = \frac{\hat{\mu}}{\partial \ln V}.$$  \hspace{1cm} (3)

If the period of an oscillating stress perturbation, $T$, is such that $T \gg t_a$, a fault system responds as if it were obeying a simple Coulomb failure model.

Fig. 8. Seasonal variations in seismicity rate and Coulomb stress rate. Top graph shows the average seasonal variation of seismicity between 04/01/1995 and 31/12/2001 (for $M_l > 3$ and $M_l > 4$) based on the declustered catalogue of (Bollinger et al., 2007). The number of events with magnitude greater than 4 has been multiplied by 5. The bottom graph shows variations in Coulomb stress rate computed at 10 km (green) and 15 km (blue) depths within the seismically active area at the front of the high range.
Our observations of seasonal correlation between seismicity and stress rate thus imply that \( t_a \) has to be significantly shorter than a year. Equation (3) then places an upper bound on \( \sigma a \) of about 8 kPa. Due to the nucleation process, the seismicity rate should not respond instantaneously to a stress rate change, but instead follow a characteristic lag time on the order of \( t_a \) (Toda et al., 2002). The small lag time between seismicity rate and stress rate seasonal changes (Fig. 7) suggests that \( t_a \) is actually significantly less than 1 yr, implying a value of \( \sigma a \) is less than 8 kPa.

When \( T < t_a \), the sensitivity of the seismicity to stress rate variations is approximated by (Dieterich, 1987; Lockner and Beeler, 1999).

\[
\frac{R_m}{r_0} = \frac{S_m}{a\sigma}
\]  

Eqs. (1) and (4) cannot be simultaneously valid except in the particular regime in which \( T \approx 2\pi t_a \). In any case, Eq. (4) always provides a lower estimate of the seismicity variations expected to result from stress variations with amplitude \( S_m \). Given that the observed seasonal variations correspond to a ratio \( \frac{R_m}{r_0} \) on the order of 1/3 and that \( S_m \) is on the order of 1–2 kPa, we infer that \( a\sigma \) is probably larger than 3–6 kPa.

These lines of reasoning both point to a value of \( a\sigma \) on the order of 3–8 kPa. This value is about one order of magnitude less than the 0.5–10 MPa range of values estimated from postseismic creep (Perfettini and Avouac, 2004; Hsu et al., 2006; Miyazaki et al., 2004; Hearn et al., 2002). Assuming a lithostatic normal stress, the 5–15 km depth range of the seismicity implies that \( a \) is on the order of \( 10^{-5} \), several orders of magnitude lower than the \( 10^{-2}–10^{-3} \) range of laboratory estimates (Marone, 1998). The low value of \( a\sigma \) might be due to high pore pressure rather than a particularly low value of \( a \). This possibility is supported by the high conductivity, suggesting the presence of fluids, observed in the area of seismicity (Fig. 2c). High pore pressure has also been proposed to explain the seasonal variation of seismicity following the Landers earthquake, which was observed mainly at sites with hydrothermal activity (Gao et al., 2000), and the small value of \( a\sigma \), about 35 kPa, which was estimated from the temporal evolution of seismicity induced by a magmatic intrusion (Toda et al., 2002).

8. Conclusion

Permanent geodetic monitoring in the Himalaya has revealed that secular interseismic strain is modulated by strong seasonal variations dominantly driven by surface load variations of hydrological origin. Himalayan background seismicity is extremely sensitive to the small seasonal stress perturbations with a peak-to-peak amplitude of a few kPa induced by this mechanism. Given a weak sensitivity to Earth tides and an insignificant lag time between stress rate and stress rate variations, we infer that earthquake nucleation in the region takes place on a time scale longer than days but probably shorter than a few months. A nucleation phase is predicted by rate-and-state friction models of earthquake nucleation, but the short nucleation inferred here requires either that \( a = \frac{\partial^2 V}{\partial r \partial t} \) on natural faults be on the order of \( 10^{-5} \)—orders of magnitude smaller than the values estimated from laboratory experiments—or, more probably in view of the correlation of seismicity with crustal conductivity, that earthquakes along the Himalayan front nucleate in areas of high pore pressure.

Acknowledgements

We are grateful to M.R. Pandey and all our collaborators, at NSC and DMG, and DASE at CEA, for their dedicated effort, which permitted the deployment, maintenance, and operation of the Nepal seismic network and the CGPS stations. This study has benefited from discussions with Hugo Perfettini, Jean-Paul Ampuero, Rodolphe Cattin, and Frederic Perrier, and from insightful reviews by Roland Burgman and Roger Bilham. Elisabeth Nadin is thanked for her help in editing this manuscript.
Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.epsl.2007.11.021.

References


