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

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 Flouzat⁴

Received 1 December 2011; revised 20 February 2012; accepted 21 February 2012; published 13 April 2012.

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° ± 3.31°, latitude = 51.4° ± 0.3° with an angular velocity of $\Omega = 0.5029 \pm 0.0072^\circ$/Myr. The pattern of coupling on the MHT is computed on a fault dipping 10° to the north and whose strike roughly 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 ± 0.4 × 10¹⁹ 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 ± 10¹⁹ 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. Introduction

The convergence rate across the central Nepal Himalaya 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 $M_W \sim 8.5$ [Ambrose 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; Seber and Gornitz, 1983].

[5] 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.

[4] 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 earthquakes [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].

[5] 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 2 and we describe in section 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 auxiliary 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 4, and then used in section 5 to evaluate the seismic hazard in Nepal. The auxiliary 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.

2. Seismotectonic Setting

[6] 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 [protect 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 very 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]. Historical large earthquakes are shown in Figure 1. 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 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.

[7] 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 advantage 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.

3. Data Used to Determine the Coupling Pattern on the MHT and the Convergence Rate

3.1. Continuous GPS Stations

[8] We use data from a network of continuous GPS stations currently consisting of 24 stations, spanning the Nepalese territory (Figure S1 in Text S1). In addition, a station in Sarangkot (SRGK) has been in operation from March 2005 to February 2007, providing one additional velocity point (see Table S2 in Text S1 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.

[9] 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 in Text 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].

[10] 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:

\[
x(t) = p_1 + p_2 t + \sum_{i=1}^{4} \left( p_{2i+1} \cos 2\pi \frac{t}{T_i} + p_{2i+2} \sin 2\pi \frac{t}{T_i} \right) + \sum_{i=1}^{S} p_{10+i} 1_{t > t_i},
\]

where \(T_i = 1\) year, \(S\) is the total number of steps known to be in the time series, \(t_i\) is the time at which each step occurs and the linear parameters \(p_i\) are defined as (1) a constant offset (parameter \(p_1\)), (2) the secular velocity (parameter \(p_2\)), (3) annual variations, modeled by sine waves of periods of 1, 1/2, 1/3 and 1/4 years (parameters \(p_3\) to \(p_{10}\)), and (4) steps in the time series (for \(S\) steps, parameters \(p_{11}\) to \(p_{10+S}\)). The function \(1_{t > t_i}\) is defined as follows:

\[
1_{t > t_i} = \begin{cases} 
0 & \text{if } t < t_i \\
1 & \text{if } t \geq t_i
\end{cases}
\]

[11] 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 sorts 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 temporally correlated and it is therefore incorrect to assume a purely white noise model [Zhang et al., 1997; Langbein and Johnson, 1997; Williams, 2003]. For these reasons we have determined a noise model and the related covariance matrix following the approach of Williams [2003] and Williams et al. [2004]. Details are given in auxiliary material.

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 Kathmandu 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

![Figure 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, and (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.](image-url)
the Garhwal-Kumaon Himalaya (India) close to the far western border of Nepal [Pouraj et al., 2011].

3.3. Determination of the Euler Pole of the Indian Plate in the ITRF2005 Reference Frame

[14] 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 Colombo. 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 auxiliary material 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 than the uncertainties 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 ± 1 mm/yr north-south shortening across the Indian subcontinent (i.e. strain rate less than 10\(^{-13}\) 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 in Figure 2 and corresponds to the following Euler pole describing the rigid Indian plate motion in the ITRF05 reference frame:

\[
\begin{align*}
\text{longitude} &= -1.34° ± 3.31° \\
\text{latitude} &= 51.4° ± 0.3° \\
\Omega &= 0.5029 ± 0.0072°/\text{Myr}
\end{align*}
\] (3)

[15] Those parameters are in good agreement with previously published GPS-based models of the Indian plate’s motion [Socquet, 2003; Bettinelli et al., 2006; Banerjee et al., 2008], as is shown in Figure 2c.

3.4. Leveling Data

[16] 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 in Figure 3a) [Jackson and Bilham, 1994]. The GPS vertical velocities have large uncertainties (Table S2 in Text S1). They are used in the inversion presented below although we find that they don’t add much constraint to the model (Figure 3b).

4. Coupling Method and Results

4.1. Inversion Method

[17] 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] doesn’t 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 3a, 5a, and 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:

\[
\text{ISC} = \frac{\text{deficit of slip rate on the patch}}{\text{long term slip rate}}. \tag{4}
\]

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.

[18] 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 in Figure 3a), 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 N98.2°E 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
Figure 3

(a) Horizontal GPS

\( V_{\text{west}} = 20.5 \pm 1.0 \text{ mm/yr} \)

\( V_{\text{east}} = 17.8 \pm 0.5 \text{ mm/yr} \)

(b) Vertical

2-GPS uplift rate

Model 30 mm/yr

Figure 3
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 very little impact on the quantities estimated in this study.

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

\[
d = Gm, \tag{5}
\]

where \(d\) is the data vector made of the surface displacements at the measurement points, \(G\) is the Green’s functions matrix computed using the semi-analytical formulation published by Okada [1985], and \(m\) 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.

[20] 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 section S3 in Text S1 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.

[21] Summarizing those constraints into an equation, the slip on the fault (vector \(m\)) is obtained by minimizing the quantity:

\[
\chi^2(m) = ||C^{-1/2}(Gm - d)||^2 + \lambda ||Am||^2 + \mu ||Im||^2, \tag{6}
\]

where \(C\) is the data covariance matrix, \(A\) is the Laplacian matrix and \(I\) is the matrix which, when multiplied by \(m\) returns the components of \(m\) for slip at depth, \(\lambda\) and \(\mu\) measure the weights attributed to each constraint. The parameter \(\mu\) 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 \(\lambda\) is chosen to minimize \(\chi^2\), which leads to \(\lambda = 0.85\) (Figure 8a). The values of the slip \(m\) are then divided by the corresponding value of the long term slip rate to obtain interseismic coupling.

[22] 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.

4.2. Results of the Inversion

[23] 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 3. The reduced chi squares of the fit to the different data sets are given in Table 1. These values show that there is no need to renormalize the uncertainties as all reduced chi-squares are of the order of unity.

[24] 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 in Figure 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.

[25] The resolution at each of the fault’s patches is plotted in Figure 5a (see section S3 in Text S1 for details on how resolution is determined). The resolution is expressed here in terms of the characteristic size of smallest inhomogeneities of coupling which could in principle be resolved given the spatial distribution and the uncertainties of the measurements. For clarity, we saturated Figure 5a at a resolution of 80 km, since one can assume that above such a value there is simply no resolution on the corresponding 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 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.

[26] 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

**Figure 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 roughly 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.
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 in Figure 6) but it could in reality be even sharper. Elsewhere the transition is always wider (light red shaded curve in Figure 6), probably because of the more limited resolution of 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.

[27] The convergence rate across the Himalaya is estimated to $17.8 \pm 0.5$ mm/yr east of the Thakkola graben and $20.5 \pm 1$ 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 χ² of the fit equal to 1. As a result, those uncertainties are slightly underestimated.

[28] The extension rate across the Thakkola graben is estimated to $3.1 \pm 2.6$ mm/yr. Elsewhere, we get a distributed extension rate of the southern Tibetan plateau of $9.3 \pm 2.1$ 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 $12.4 \pm 4.7$ nstrain. yr⁻¹, consistent with, although somewhat lower than the $\sim 35$ 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 $18.1 \pm 0.5$ mm/yr east of the Thakkola graben and $20.8 \pm 1$ 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%

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Reduced χ²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Continuous GPS</td>
<td>1.92</td>
</tr>
<tr>
<td>Campaign GPS</td>
<td>2.94</td>
</tr>
<tr>
<td>Leveling</td>
<td>1.69</td>
</tr>
<tr>
<td>Total</td>
<td>2.29</td>
</tr>
</tbody>
</table>

Table 1. Values of the Reduced χ² of the Fit to the Different Data Sets

Figure 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.
Figure 5. (a) Resolution on each of the patches of the fault, given in km. See text and auxiliary material for details on the computation of the resolution. Locations of the data points used to compute the resolution are indicated. (b) Map of the coupling on the MHT on which patches with no resolution (resolution > 80 km) are masked.
different from the ones derived with the model of linear arc parallel extension in southern Tibet. Figure S5 in Text S1 shows that the direction selected for the arc-parallel extension doesn’t have an significant impact either.

5. Implications

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 ± 2.5 mm/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.  
\[30\] 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].

5.2. Temperature Control on the Downdip End of the Locked Fault Zone

\[31\] 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 5). Figure 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 in Figure 6 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 quartzofeldspathic 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 advocated as an explanation for the seismicity cut-off generally observed at a depth of around 15 km 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.

5.3. Relationship Between Geodetic Strain and Background Seismicity

\[32\] 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 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 \(20 \times (1-\text{ISC})\) 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.

\[33\] 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 in Figure 7 should be seen as lower boundaries of the real values.

\[34\] In fact, the seismicity does not occur on the MHT per se but rather within a 5–10 kilometer size volume around the downdip end of the locked fault zone. The moment released by the background seismicity amounts to \(1.2 \times 10^{17}\) Nm/yr which represents less than 0.2% of the deficit of moment accumulating due to interseismic locking of the MHT (Table 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 \(M_L^{\text{NSC}}\) reported in the NSC catalogue were converted into moment magnitudes \(M_w^{\text{NSC}}\) from a subset of events reported also in the CMT catalogue (L. Bollinger, NSC, personal communication, 2011) yielding:

\[7\]

\[M_w^{\text{NSC}} = 0.84M_L^{\text{NSC}} + 0.21.\]

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

\[35\] 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 3500 m (Figure 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 3500 m 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].

5.4. Moment Deficit Accumulation Rate, Return Period and Magnitude of the Largest Plausible Earthquake

\[36\] 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 $\dot{M}_0 = 6.6 \pm 0.4 \times 10^{19}$ 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 8a) or the direction selected for the extension of the Tibetan plateau (Figure S5 in Text S1).

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 in Figure S3 in Text S1), 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 roughly match the geological slip rate 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 doesn’t 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

### Table 2. Moment Released During Earthquakes Annually According to Different Seismicity Catalogs

<table>
<thead>
<tr>
<th>Catalog</th>
<th>Time Span (yrs)</th>
<th>Moment Released Seismically (Nm/yr)</th>
<th>Fraction of Moment Accumulated (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Historic</td>
<td>~500</td>
<td>$1.8 \times 10^{19}$</td>
<td>27.3</td>
</tr>
<tr>
<td>Historic</td>
<td>~500</td>
<td>$0.9 \times 10^{19}$</td>
<td>13.6</td>
</tr>
<tr>
<td>CMT</td>
<td>35</td>
<td>$2.9 \times 10^{18}$</td>
<td>4.4</td>
</tr>
<tr>
<td>CMT micro</td>
<td>35</td>
<td>$3.4 \times 10^{18}$</td>
<td>0.5</td>
</tr>
<tr>
<td>NSC</td>
<td>6</td>
<td>$1.2 \times 10^{17}$</td>
<td>0.2</td>
</tr>
</tbody>
</table>

*This moment released is compared to the rate of moment deficit of $6.6 \times 10^{19}$ Nm/yr inferred in this study.

*The value considered for the total moment deficit accumulated is the one derived in this study: $\dot{M}_0 = 6.6 \pm 0.4 \times 10^{19}$ Nm/yr (see section 5.4).

*Catalog of earthquakes happening on the MHT extracted from Ambraseys and Douglas [2004], except for the 1950 Assam earthquake, for which the instrumental moment magnitude computed by Chen and Molnar [1977] was used.

*Catalog of earthquakes happening on the MHT extracted from Ambraseys and Douglas [2004] only.

*CMT catalog without considering the 2005 Kashmir earthquake and its aftershocks.

*See text for details on how the NSC moment magnitude is evaluated.
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.

[39] We also assume that a fraction $\alpha$ 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 $M_{\text{max}}$, 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.

[40] 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 Figure S3b in Text S1). 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 couldn’t 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 don’t 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., 2009; 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 $\alpha$ 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 $\alpha$ value ranging between 0.6 and 0.7 [Jouanne et al., 2011]. So a reasonable range of values for $\alpha$ is probably between 0.5 and 0.9, with a more probable value around 0.8.

[41] Under those assumptions the recurrence time of earthquakes of moment $M$ is [Molnar, 1979]

$$T(M) = \frac{1}{1 - 2b/3} \frac{\alpha M_{\text{max}}}{\alpha M_0} \left( \frac{M}{M_{\text{max}}} \right)^{2b/3},$$

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 8b) the return period of the largest possible earthquakes on the MHT becomes

$$T_{b=1}(M_{\text{max}}) = \frac{3M_{\text{max}}}{\alpha M_0}.$$  

[42] Figure 8b displays a comparative Gutenberg-Richter plot of the different seismicity catalogs available in Nepal and equation (9) for $\alpha = 1$, and $M_{\text{max}} = 8, 9$ 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.

[43] 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° ± 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 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 (1905, 1934, 1950), and in the past 200 years, $6 M_W \geq 7.5$ earthquakes.

[44] Above their magnitude of completeness, the earthquakes catalogs are fairly consistent with $b = 1$. However, they fall noticeably below the 3 dashed lines corresponding to equation (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 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 $M_W^{\text{max}} = 9.2$ would have a return period of the order of 3000 years if all the moment deficit was released seismically (see Figure S6a in Text S1). 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 hasn’t 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.

6. Conclusion

[45] 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 100 km. Interseismic coupling decreases abruptly, within a transition zone probably narrower than 30 km. 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., 2009]. This favors the scenario of a primary control by temperature of the locked-creeping transition, similar to the conclusions drawn by Hsu et al. [2009] on the Chelungpu fault in Taiwan and by Brooks et al. [2011] on the Mandeyapecau thrust fault in the...
Acknowledgments. This project was supported by the Gordon and Betty Moore Foundation, through the Tectonics Observatory, NSF grant EAR0834749 and the China Geological Survey (121201112165). 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 193.

References


Kumar, S. S. G. Wesnosky, 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


T. Ader, J.-P. Avouac, K. Chanard, J. Galetzkia, J. Genrich, and M. Thomas, Department of Geological and Planetary Sciences, California institute of Technology, MC 252-21, Pasadena, CA 91125, USA. (ader@caltech.edu)

L. Bollinger and M. Flouzat, Commissariat à l’Energie Atomique, DAM, DIF, F-91297 Arpajon CEDEX, France.

L. Ding and J. Liu-Zeng, Key Laboratory of Continental Collision and Tibetan Plateau Uplift, Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing 100085, China.

H. Lyon-Caen, Laboratoire de Géologie, École Normale Supérieure, L. Ding and J. Liu-Zeng, Key Laboratory of Continental Collision and Tibetan Plateau Uplift, Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing 100085, China.

S. Rajaure, S. N. Sapkota, and P. Shrestha, National Seismological Centre, Department of Mines and Geology, Lainchaur, Kathmandu, Nepal.