Earth and Planetary Science Letters ••• (••••) •••-•••



Contents lists available at SciVerse ScienceDirect

Earth and Planetary Science Letters



EPSL:12087

www.elsevier.com/locate/epsl

# Detecting periodicities and declustering in earthquake catalogs using the Schuster spectrum, application to Himalayan seismicity

# Thomas J. Ader<sup>a,b,\*</sup>, Jean-Philippe Avouac<sup>a</sup>

<sup>a</sup> Department of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA 91125, USA
 <sup>b</sup> Laboratoire de Géologie, Ecole Normale Supérieure, CNRS, 24 rue Lhomond, 75004 Paris, France

#### ARTICLE INFO

Article history: Received 18 October 2012 Received in revised form 13 June 2013 Accepted 23 June 2013 Available online xxxx Editor: Dr. P. Shearer

Keywords: earthquake catalogs periodicities Schuster test Schuster spectrum Nepalese seismicity seasonal variations

## ABSTRACT

We show that the Schuster test alone does not provide a sufficient condition to assert the existence of a periodicity in an earthquake 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). We show that the detection level is slightly period dependent, and we provide an analytical expression relating the amplitude of seismicity-rate variations to the confidence level at which the probability that the observed variations be due to chance can be discarded. The Schuster spectrum also provides information about the deviation from a sinusoidal function of the periodicity of the seismicity rate, and identifies an eventual imperfect declustering of the catalog, making it coincidently a potential tool to determine whether or not a catalog has been properly declustered. Applying this tool to the Nepalese seismicity appears. In particular, no variations of seismicity rate of amplitude up to 40%, while no other periodicity appears. In particular, no variations of seismicity at these periods is less than 18%.

Published by Elsevier B.V.

### 1. Introduction

A number of studies have investigated earthquake mechanics by examining the response of seismicity to periodic stresses. The investigation is regularly carried out under the assistance of the Schuster test (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 quantitatively refute the claim by Knott (1897) that earthquakes and tides were correlated in Japan. Tanaka et al. (2002a, 2006) and Cochran et al. (2004) revisited recently 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 the probability to discard 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 forcing 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

E-mail address: ader@caltech.edu (T.J. Ader).

0012-821X/\$ - see front matter Published by Elsevier B.V. http://dx.doi.org/10.1016/j.epsl.2013.06.032 annual variations of microseismicity in Nepal with it, which they related 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 seismicity variations related to a periodic perturbation.

In this study, we call for caution when using the Schuster test, by pointing out that while a periodicity in an earthquake catalogue may yield a low Schuster *p*-value, the reverse is not necessarily true: a low Schuster *p*-value may not mean that the seismicity rate is periodic at the period considered. To circumvent this ambiguity, we propose an extension of the Schuster test, providing a way to assert the existence of a periodicity in an earthquake catalog, which may also conversely be used to identify periodicities of the seismicity rate in the catalog at unexpected periods. The idea is to compute a spectrum of Schuster *p*-values at a proper subset of periods within a given range, hereafter referred to as the Schuster spectrum.

After briefly presenting the principle of the Schuster test, Section 3 determines the appropriate method to build the Schuster spectrum, as well as the confidence level at 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 4, emphasizing the fact that the claim of periodic variations of the seismicity rate in a catalog requires

<sup>\*</sup> Corresponding author at: Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA 91125, USA.

T.J. Ader, J.-P. Avouac / Earth and Planetary Science Letters ••• (••••) •••-•••

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 5, we show that annual variations of the seismicity rate are discernible for the largest events of the catalogs, whereas they cannot be unequivocally claimed for smaller events. These annual variations are the only ones that come out of the spectra. In particular, we show that no variations at the tidal periods are apparent, and derive quantitative constraints on the maximum amplitude of the possible variations of the seismicity rate at tidal periods.

### 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 sinusoidal variations of the seismicity rate. To compute the probability that the timing of events in a catalog varies according to a sine-wave function of period T, a phase is associated to each event. Calling  $t_k$  the time of event number k, its associated phase  $\theta_k$  is:

$$\theta_k = 2\pi \frac{t_k}{T}.\tag{1}$$

The catalog of times can hence be converted into a 2D walk, made of successive unit-length steps, in directions given by these phases. Denoting by D the distance between the start and end points of this walk, the probability p, that a distance greater than or equal to D 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):

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

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 that the distribution of the timings of events stacked over the period T is non-uniform, which is usually interpreted as the probability of a periodicity at period T.

If a catalog contains *N* events occurring out of a sinusoidallyvarying seismicity rate:

$$\frac{R(t_k)}{r} = 1 + \alpha \cos\left(\frac{2\pi t_k}{T}\right),\tag{3}$$

where *r* is the average seismicity rate,  $\alpha$  is the amplitude of the seismicity-rate variations,  $t_k$  is the time of event number *k*, and *T* is the period of the variations, the logarithm of the Schuster *p*-value computed at period *T* follows (see supplementary material, Section S.1 for derivation):

$$\begin{cases} \langle -\ln p \rangle = \left(\frac{D^2}{N}\right) = 1 + \frac{N\alpha^2}{4}, \\ \operatorname{var}(\ln p) = \operatorname{var}\left(\frac{D^2}{N}\right) = \left(1 - \frac{\alpha^2}{2}\right) \left(1 + \frac{N\alpha^2}{2}\right). \end{cases}$$
(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  $\alpha$  of the seismicity-rate variations.

## 3. Building a spectrum of Schuster p-values

### 3.1. Schuster spectrum vs isolated Schuster test

The process of testing a periodicity at period T in an earthquake catalog always boils down to the same underpinnings: the



**Fig. 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: sinusoidal 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 during the stacking process. It is therefore redundant to test periodicities both at *T* and  $T + \Delta T_1$ , but not at *T* and  $T + \Delta T_2$ .

catalog gets stacked over the investigated period T, and the probability that the stacked catalog is non-uniform is subsequently estimated with whatever test is chosen (the Schuster test, fitting a sine-wave function to the stacked catalog, etc.). Therefore, any non-uniformity of the seismicity rate that is preserved throughout the stacking process might result in a low Schuster p-value. For instance, both a seismicity rate with periodic variations at period Tor a sudden burst of seismicity of duration less than T may yield a low *p*-value, indicating that a low Schuster *p*-value cannot alone indicate unambiguously a periodicity of the seismicity rate. As we show in the rest of this section and in Section 4, performing a set of Schuster tests over a continuous range of periods enables one to resolve this issue. The result of performing such a set of tests, i.e., computing a Schuster spectrum, is twofold: it provides a way to assert the existence of an expected periodicity, and a way to detect any unknown periodicity in a catalog. In the following, we first determine the subset of periods at which the Schuster test has to be performed in order to build a comprehensive spectrum. We then establish the confidence levels for the detection of periodicities in the so-built Schuster spectrum.

### 3.2. Subset of periods for the spectrum

In order to determine the finite subset of periods to select in order to test all the periods within a continuous range, one should note that when testing a period T with the Schuster test, the continuous range of periodicities around T that remain coherent throughout the stacking process gets tested. Fig. 1 illustrates this point. Let us 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. Instead of testing the period T, when 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), which 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.

T.J. Ader, J.-P. Avouac / Earth and Planetary Science Letters ••• (••••) •••-•••



EPSL:12087

3

**Fig. 2.** 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 Eq. (9), while "99% confidence level" corresponds to 1% of minimum expected values.

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:

$$n(T_i)\Delta T_i = \varepsilon T_i,\tag{5}$$

where  $\varepsilon$  will be determined more precisely later. Noting that  $n(T_i) = \mathcal{I}(t/T_i) \leq t/T_i$ , the condition in Eq. (5) can be replaced by the following condition for the period increment:

$$\Delta T_i = \frac{\varepsilon T_i^2}{t}.\tag{6}$$

Noting v = 1/T the frequency, the frequency increment is thus constant:

$$\Delta \nu = \frac{\varepsilon}{t},\tag{7}$$

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).

Eq. (6) shows that the period increment is smaller at short periods than at large periods. For instance, taking  $\varepsilon = 1$ , for an earthquake catalog of duration t = 10 yr, the period increment has to be  $\Delta T_{\text{tides}} \approx 7 \times 10^{-5}$  days ( $\approx 6$  s) in order to detect any periodicity around the main tides period (i.e.,  $T_{\text{tides}} \approx 0.5$  days), 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  $v_{\min} = 1/T_{\max}$  and  $v_{\max} = 1/T_{\min}$ ) requires performing *N* Schuster tests, where from Eq. (7):

$$N = \frac{t}{\varepsilon} \left( \frac{1}{T_{\min}} - \frac{1}{T_{\max}} \right) \approx \frac{t}{\varepsilon T_{\min}},\tag{8}$$

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

As is suggested by Eq. (6), the subset of periods at which the spectrum will be computed depends on the choice of  $\varepsilon$ . 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. The optimal value of  $\varepsilon$  is determined in the supplementary material, Section S.2, and is  $\varepsilon_0 \approx 1$ . For safety, the results might be double-checked by oversampling the spectrum, although  $\varepsilon$ -values significantly smaller than 1 do not bring any additional information and require longer computation times. In practice, our experience is that  $\varepsilon$ -values smaller than 0.5 are certainly not useful.

#### 3.3. Confidence levels for the detection of periodicities

With the value of  $\varepsilon_o$  in hand, it is possible to estimate the expected threshold above which a Schuster *p*-value in the spectrum indicates with confidence that the seismicity rate contains a periodicity. As is explained in the supplementary material, Section S.2, the Schuster *p*-value for a catalog occurring out of a uniform seismicity rate is uniformly distributed over the interval [0; 1]. As a result, the expected value of the minimum Schuster *p*-value in a spectrum containing *N* independent tests 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 Schuster probability for periods greater than *T* is simply:

$$\langle \delta_m \rangle = \frac{T}{t}.\tag{9}$$

A periodicity in the catalog will thus have a significant probability not to be due to chance if its Schuster *p*-value is significantly lower than this expected value. Quantitatively, a periodicity can be 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. Fig. 2 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, consistent with Eq. (9).

As suggested by Eq. (7), a "flat" spectrum would be obtained for a linear *x*-axis in frequencies. With such a representation, the expected Schuster *p*-values line would not be 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 at shorter periods. However, if Eq. (9) may suggest a drastic dependence of the detection threshold on the period, combining it with Eq. (4) leads to the following expression for the critical amplitude of seismicity rate variations necessary for a detection at the 95% confidence level:

$$\alpha_{95} = \frac{2}{\sqrt{N}} \sqrt{2 + \ln \frac{t}{T}}.$$
(10)

Eq. (10) indicates that the critical amplitude of the seismicity-rate variations above which a periodicity can generally be detected is only slightly sensitive to the period. For a 1000-event catalog covering 10 yr,  $\alpha_{95}(T = 1 \text{ yr}) \approx 14\%$ , while  $\alpha_{95}(T = T_{\text{tides}}) \approx 21\%$ .



**Fig. 3.** (a) Schuster spectrum for a 1000-event catalog generated out of a sinusoidal seismicity rate following Eq. (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-sinusoidal seismicity rate following Eq. (11), with variations of amplitude  $\alpha = 3$ , and period T/t = 0.04. (c) Schuster spectrum for a 1000-event catalog containing an aftershock sequence, which seismicity rate is given by Eq. (12), and 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.

# 4. Application to synthetic catalogs

Now that we have exposed how to build a Schuster spectrum and established the levels of confidence for the detection of periodicities, we apply it to three different synthetic catalogs. We show that it allows detecting an unknown periodicity, and that it actually is the only way to assess whether the catalog analyzed contains a periodicity or not. Indeed, if variations of the seismicity rate remain coherent throughout the stacking process, the Schuster *p*-values are small, whether or not the variations of seismicity rate are periodic at the period *T*. In particular, periodic variations of the seismicity rate at periods that are an integer multiple of *T*, or sudden outbursts of seismicity, will lead to low *p*-values. We consider the three following types of catalogs: one generated out of a sinusoidal seismicity rate, one out of a periodic but non-sinusoidal seismicity rate, and one out of a uniform seismicity rate superimposed with an aftershock sequence. Supplementary material Section S.4 describes in details how the catalogs have been generated.

First of all, Fig. 3(a) represents the Schuster spectrum from a 1000-event catalog generated out of a sinusoidal seismicity rate following Eq. (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 not sinusoidal, harmonics of the main period may appear in the spectrum.

5

Fig. 3(b) shows the spectrum for a catalog generated out of the following periodic seismicity rate:

$$\frac{R(t_k)}{r} = \begin{cases} 1, & \text{if } t_k[T]/T \in [0; 0.1], \\ \alpha, & \text{if } t_k[T]/T \in [0.1; 0.2], \\ 1, & \text{if } t_k[T]/T \in [0.2; 1], \end{cases}$$
(11)

where  $t_k$  is the time of event number k, T is the period of the seismicity-rate variations,  $t_k[T]$  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 Fig. 3(b). The seismicity rate used to generate the catalog analyzed in Fig. 3(b) was obtained for T/t = 0.04 and  $\alpha = 3$ . In this case, the periodicity at T/t = 0.04 appears clearly, but harmonic periodicities can also be noticed 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, although the smallest harmonic appearing in the spectrum provides an estimate of the duration of the higher seismicity rate within one period.

Another important 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 in Fig. 3(c): the simulated catalog has a uniform background seismicity rate is described using Dieterich (1994) aftershocks model:

$$\frac{R(t_k)}{r} = \frac{1}{1 + (e^{-Q} - 1)e^{-(t_k - t_m)/t_a} \mathcal{H}(t_k - t_m)},$$
(12)

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  $x \ge 0$ ). We suppose that  $t_a \ll t - t_m$  in order to make sure that the aftershock sequence is over before the end of the catalog. The catalog used to generate Fig. 3(c) contains a background of 1000 events and an aftershock sequence containing 100 events, with  $t_a/t = 10^{-2}$ ,  $t_m/t = 0.2$  and Q = 10. In this case, for all periods of the order of or larger than the characteristic duration of the evolution of the seismicity rate in the aftershock cluster, the "Schuster walk" progresses 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. Note that the same type of spectrum will result from the presence of earthquake swarms in a catalog or any combination of increases of the seismicity rate over durations 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 event 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, since in this case, the spectrum contains many peaks above the 95% confidence level for periods above a given period. 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 an isolated Schuster test. It provides an unambiguous diagnostic on whether a catalog contains a periodicity, or if low Schuster *p*-values are due to different non-uniformities of the seismicity rate.

#### 5. Application to the seismicity of Nepal

A large fraction of the Nepalese seismicity clusters along the down-dip end of the locked part of the Main Himalayan Thrust fault (MHT) (Cattin and Avouac, 2000; Bollinger et al., 2004; Ader et al., 2012). Most of the activity comes from thrust events induced by north-south compression related to the ongoing convergence across the Himalaya, 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. Analysing 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, although the test was only performed at the period T = 1 yr. This seasonal modulation of the seismicity rate is attributed to surface load variations following the hydrological cycle, which induce fluctuations of the Coulomb stress in the zone of seismicity, with a peak-to-peak amplitude of 2 to 4 kPa, whereas the interseismic stress builds up within the seismicity cluster at a rate of 10 to 20 kPa/yr (Bollinger et al., 2004, 2007; Bettinelli et al., 2008). Bettinelli et al. (2008) also noted that, although the variations of stress due to solid-Earth tides are of similar amplitude as those caused by the hydrological loading, the Nepalese seismicity does not seem to correlate with tides. However, no constraints on the possible maximum amplitude of variations at the tidal periods were determined.

The Schuster spectrum applied to the NSC seismicity catalog, which now extends until the end of 2008, allows placing quantitative constraints on the amplitude of the variations of the seismicity rate on the MHT at the annual and tidal periods. In parallel, we corroborate our results by also analyzing the ISC catalog (International Seismological Centre, 2010) from 1965 to 2008. Events from both the ISC and NSC catalogs in the midcrustal cluster are selected using the same contour in map view as in Bollinger et al. (2007). Both catalogs are declustered with the algorithm described in Reasenberg (1985), with the same set of parameters as in Bollinger et al. (2007) (P = 0.95,  $1 \le \tau \le 10$  days,  $D \le 20$  km,  $U_x = 5$  km and  $U_z = 10$  km). Fig. 4 shows the temporal evolution of  $M_L \ge 3$  selected events from both the raw (i.e., not declustered) 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, Fig. S3).

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, is 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  $M_L \ge 3$ , which is above the detection level at all time (Bollinger et al., 2007). The Schuster spectrum of the non-declustered catalog used in the study by Bollinger et al. (2007) (Fig. 5(a)) displays a prominent peak at 1 yr, but it also contains numerous peaks at larger and smaller periods,



**Fig. 4.** Time and space distribution of the  $M_L \ge 3$  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  $M_L \ge 5.5$  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). Circle sizes are proportional to event magnitudes: smallest events have  $M_L = 3$ , while  $M_L \ge 5.5$  events are in blue and their magnitude is indicated, giving an idea of the scale. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



**Fig. 5.** Schuster spectra built with the NSC  $M_L \ge 3$  seismicity catalog (a) over the same period of time as in Bollinger et al. (2007) (i.e. from 1995 to 2001) and without declustering, and (b) for the whole available NSC  $M_L \ge 3$  declustered catalog (i.e. from 1995 to 2008). The blue dashed lines indicate tidal, half annual and annual periods. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

T.J. Ader, J.-P. Avouac / Earth and Planetary Science Letters ••• (••••) •••-•••



**Fig. 6.** 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  $M_L \ge 5.5$  events for the entire available catalog (1995 to 2008). The spectrum for the ISC catalog is computed for  $M_b \ge 4$  events from 1965 to 2008. The annual peak is highlighted in red for clarity. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

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 (Fig. 4). 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 Fig. S4, supplementary material), that remains even after declustering, and which would induce low p-values at periods larger than about a year. Once the NSC catalog is declustered and extended until the end of 2008, most of the previous periodicities disappear from the spectrum (Fig. 5(b)), but peaks at periods larger than 2.5 yr consistently remain, suggesting that clusters are most likely still present in the catalog, and were not properly removed by the declustering. For periods around 1 yr and above, the Schuster p-values therefore cannot be interpreted in terms of periodicities in the catalog. The same conclusion persists when looking at the declustered catalog from Bollinger et al. (2007) (supplementary Fig. S5). 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, such as larger seismic events, slow slip events, sub-surface hydrology, mining, etc. The Schuster spectrum for NSC  $M_L \ge 3$  events thus shows that no rigorous claim can be made for periods of 1 yr or more because of clusters present in the catalogs, which characteristic duration is less than 1 yr.

However, as the application to synthetic catalogs shows (Fig. 3(c)), for periods much less than the typical duration of clusters of seismicity, the spectrum is not affected by the clusters. In the case of the 1995–2008 declustered NSC catalog, the Schuster spectrum at the diurnal and semi-diurnal tidal periods does not seem affected by the clusters still present in the catalog (Fig. 5). At these periods, the fact that no Schuster *p*-value exceeds the 95% confidence level is not an artifact of an improper declustering of

the catalog. The lowest Schuster *p*-value in this region of the spectrum is  $10^{-4.5}$ , which corresponds to an amplitude  $\alpha = 15\% \pm 3\%$ , according to the system of Eqs. (4). At these periods, the relative variations of the seismicity rate  $\alpha$  must therefore be less than 18%.

In order to circumvent the declustering issue at the annual period, we look at events of larger magnitude, which tend to cluster less. This is now possible thanks to the longer time span of the NSC declustered catalog, which contains 16 events of  $M_L \ge 5.5$ 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 in Fig. 4), a misbalance that only has a 2% binomial probability to happen out of a uniform seismicity rate. Some of the NSC  $M_L \ge 5.5$  events may appear to cluster in time in Fig. 4, but a longitude-time plot of the seismicity shows that these events are sufficiently far apart in space that they are not related to each other (supplementary Fig. S8). On the Schuster spectrum computed for these events (Fig. 6(a)), the annual periodicity appears as the only one above the 95% confidence level, suggesting indeed that peaks at other periods on the spectrum for  $M_L \ge 3$  events were mostly due to isolated variations of the seismicity rate rather than true periodicities. Besides, here again, no variations at the tidal period stand out.

The Schuster spectrum computed over the 210  $M_b \ge 4$  events from the declustered ISC catalog taken from 1965 to 2008 (Fig. 6(b)) backs up these observations: the periodicity at 1 yr still emerges alone above the 95% detection level. Eq. (4) indicates that the Schuster *p*-value at 1 yr corresponds to seismicity-rate variations of amplitude  $\alpha = 27\% \pm 7\%$ , 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 the seismicity rate thus prevail for

EPSL:12087

7

T.J. Ader, J.-P. Avouac / Earth and Planetary Science Letters ••• (••••) •••-••



**Fig. 7.** Evolution of the Schuster *p*-value by adding years to the ISC declustered catalog, considering  $M_b \ge 4$  events, starting in 1965. The dashed black line indicates 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 Eq. (4).

larger events, and stand alone as the only periodic variations of the midcrustal seismicity in Nepal.

Using Eqs. (4) and (9), together with current seismicity rates from the NSC catalog, indicates that 6 yr of the declustered  $M_L \ge 3$ NSC catalog should be enough to detect variations of the seismicity rate of amplitude  $\alpha$  greater than 20% above the 95% confidence level, and that the 14 available years should actually be sufficient to detect variations of amplitude as low as 15% (or 22% when looking at  $M_L \ge 3.5$  events). Supplementary Fig. S6 more comprehensively indicates the minimum theoretical catalog duration necessary to detect annual variations of seismicity at the 95% confidence level, for different cutoff magnitudes of the NSC catalog. Were the annual variations of the seismicity rate of the  $M_L \ge 3$ NSC events as intense as those of the  $M_b \ge 4$  ISC events (i.e., 27%)  $\pm$  7%), they should thus clearly come out in the spectra. But as showed earlier, once the NSC catalog is declustered, annual variations of the seismicity rate might appear when considering the catalog up to 2001 (6 yr of data, Fig. 5(a)), although they cannot be told apart from the presence of clusters of seismicity, and they clearly do not show up anymore in the 1995-2008 catalog (Fig. 5(b)).

As has been discussed before, an incomplete declustering of the catalog might be responsible for biased Schuster *p*-values at the annual period, interfering either constructively or destructively with the periodicity at 1 yr. Moreover, if the seismicity rates of both larger and smaller events follow similar periodic trends, the declustering process might smear out the seismicity-rate variations of smaller events, as smaller events resulting from the periodic seismicity-rate increase might be mistaken for aftershocks of the larger events, and thereupon removed from the catalog through declustering. This would explain in retrospect why annual variations of seismicity appear so clearly in the undeclustered catalog (Fig. 5(a)), although, once again, this is rigorously not possible to assert it with this spectrum. More than a data-processing artifact, the neutralization of seasonal variations of smaller events by the seasonal variations of the rate of larger events could also be a real process. Large winter events may in fact trigger surrounding faults close to failure as part as their aftershock sequences, whereas these faults would have otherwise ruptured later in the winter, following the gradual increase of seismicity rate subsequent to the periodic variations. 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 the rate of occurrence of independent events.

It is also possible that the amplitude of the seasonal variations varies with time, which could be due to an actual change of the response of Nepalese seismicity to seasonal perturbations of stress, but could also be a simple statistical illusion. The oc-

currence of a large event has been inferred to be able to impact the sensitivity of microseismicity to periodic stress perturbations (e.g., Tanaka et al., 2002b). More generally, any gradual change of faults properties in Nepal could result in a change of the sensitivity on the local seismicity to annual stress perturbations. However, the change of amplitude of the response of Nepalese seismicity to seasonal perturbations of stress remains within its expected statistical deviation, given the number of events considered. Fig. 7 shows the evolution of the Schuster *p*-value at the 1-yr period when years are successively added to the  $M_b \ge 4$  declustered ISC catalog. When compared to the expected evolution of the Schuster p-value, yielded by the system of Eqs. (4) with variations of amplitude  $\alpha = 40\%$ , the computed *p*-values remain within one standard deviation from the expected values. Note that the value  $\alpha = 27\% \pm 7\%$  specifically corresponds to the Schuster *p*-value considering the entire 1965-2008 catalog (rightmost value of the black curve in Fig. 7), whereas the value  $\alpha = 40\%$  seems to better follow the gradual trend of evolution of the annual Schuster p-value in Fig. 7 throughout this time span. This plot therefore indicates that the decrease of measured  $\alpha$  after 2000 might simply be a statistical effect rather than a true loss of the seasonal modulation of the seismicity rate. The amplitude of annual variations of the seismicity rate may globally be around 40%, but may appear to be less when looking at a shorter time range. This decrease seems to also appear in the  $M_L \ge 3$  declustered NSC catalog (supplementary Fig. S7), although Fig. S7 can be misleading since, as has been explained earlier, the Schuster *p*-value at 1 yr is also affected by isolated and clustered variations of the seismicity rate at such low magnitudes.

### 6. Conclusion

We propose a way to extend the Schuster test by building an entire spectrum of Schuster *p*-values. The obtained spectrum provides an efficient tool to detect unknown periodicities in an earthquake catalog, and also to determine whether or not variations of the seismicity rate in a catalog are actually periodic, which cannot be stated with confidence by an isolated Schuster test.

Applying this Schuster spectrum to earthquake catalogs from the midcrustal cluster of seismicity in Nepal suggests that intermediate events ( $M_L \ge 5.5$  or  $M_b \ge 4$ ) exhibit seasonal variations of seismicity, with an increase of seismicity in the winter as high as 40%. 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 rate. No other periodic variations of seismicity rate ap-

q

pear in the catalog. In particular, no periodicity at any of the tidal periods is detected, indicating that variations of the seismicity rate at tidal periods, if they exist, cannot be of amplitude greater than 18%.

The implementation of the spectrum is straightforward. Our implementation, written in Matlab, can be found on the Tectonics Observatory's website (http://www.tectonics.caltech.edu/ resources). The code Schuster\_test\_log.m computes the logarithm of the *p*-values for a given catalog at a given array of periods, while the code Schuster\_spectrum.m computes and plots the whole Schuster spectrum of a catalog between two given periods.

# Acknowledgements

We thank Francisco Ortega for his help in writing some of the algorithms for the simulations, run on the parallel computer Fram available at Caltech. 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 230 and Caltech Seismolab contribution number 10092.

## Appendix A. Supplementary material

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.epsl.2013.06.032.

### References

- Ader, T., Avouac, J.-P., Liu-Zeng, J., Lyon-Caen, H., Bollinger, L., Galetzka, J., Genrich, J., Thomas, M., Chanard, K., Sapkota, S.N., Rajaure, S., Shrestha, P., Ding, L., Flouzat, M., 2012. Convergence rate across the Nepal Himalaya and interseismic coupling on the Main Himalayan Thrust: Implications for seismic hazard. J. Geophys. Res., Solid Earth 117, B04403.
- Beeler, N., Lockner, D., 2003. Why earthquakes correlate weakly with the solid Earth tides: Effects of periodic stress on the rate and probability of earthquake occurrence. J. Geophys. Res., Solid Earth 108 (B8).
- Bettinelli, P., Avouac, J.-P., Flouzat, M., Bollinger, L., Ramillien, G., Rajaure, S., Sapkota, S., 2008. Seasonal variations of seismicity and geodetic strain in the Himalaya induced by surface hydrology. Earth Planet. Sci. Lett. 266, 332–344.

- Bollinger, L., Avouac, J., Cattin, R., Pandey, M., 2004. Stress buildup in the Himalaya. J. Geophys. Res., Solid Earth 109 (B11), B11405.
- Bollinger, L., Perrier, F., Avouac, J.-P., Sapkota, S., Gautam, U., Tiwari, D.R., 2007. Seasonal modulation of seismicity in the Himalaya of Nepal. Geophys. Res. Lett. 34, L08304.
- Burtin, A., Bollinger, L., Vergne, J., Cattin, R., Nábelek, J.L., 2008. Spectral analysis of seismic noise induced by rivers: A new tool to monitor spatiotemporal changes in stream hydrodynamics. J. Geophys. Res. 113, B05301.
- Cattin, R., Avouac, J., 2000. Modeling mountain building and the seismic cycle in the Himalaya of Nepal. J. Geophys. Res., Solid Earth 105, 13389–13407.
- Cochran, E.S., Vidale, J.E., Tanaka, S., 2004. Earth tides can trigger shallow thrust fault earthquakes. Science 306, 1164–1166.
- Dieterich, J.H., 1994. A constitutive law for rate of earthquakes production and its application to earthquake clustering. J. Geophys. Res. 99 (B2), 2601–2618.
- Heaton, T.H., 1975. Tidal triggering of earthquakes. Geophys. J. R. Astron. Soc. 43, 307–326.
- Hernandez, G., 1999. Time series, periodograms, and significance. J. Geophys. Res. 104 (A5), 10355–10368.
- International Seismological Centre, 2010. On-line Bulletin. Int. Seis. Cent., Thatcham, UK. http://www.isc.ac.uk.
- Knott, C., 1897. On lunar periodicities in earthquake frequency. Proc. R. Soc. Lond. 60, 457–466.
- Lockner, D., Beeler, N., 1999. Premonitory slip and tidal triggering of earthquakes. J. Geophys. Res., Solid Earth 104 (B9), 20133–20151.
- Pandey, M., Tankudar, R., Avouac, J.-P., Lavé, J., Massot, J.-P., 1995. Interseismic strain accumulation on the Himalayan crustal ramp (Nepal). Geophys. Res. Lett. 22, 751–754.
- Rayleigh, L., 1880. On the result of a large number of vibrations of the same pitch and arbitrary phase. Philos. Mag. 10, 73–78.
- Reasenberg, P., 1985. Second-order moment of Central California seismicity, 1969–1982. J. Geophys. Res. 90, 5479–5495.
- Rydelek, P.A., Hass, L., 1994. On estimating the amount of blasts in seismic catalogs with Schuster's method. Bull. Seismol. Soc. Am. 84 (4), 1256–1259.
- Scargle, J.D., 1982. Studies in astronomical time series analysis. II. Statistical aspects of spectral analysis of unevenly spaced data. Astrophys. J. 263, 835–853.
- Schuster, A., 1897. On lunar and solar periodicities of earthquakes. Proc. R. Soc. Lond. 61, 455–465.
- Tanaka, S., Ohtake, M., Sato, H., 2002a. Evidence for tidal triggering of earthquakes as revealed from statistical analysis of global data. J. Geophys. Res. 107, 2211.
- Tanaka, S., Ohtake, M., Sato, H., 2002b. Spatio-temporal variation of the tidal triggering effect on earthquake occurrence associated with the 1982 South Tonga earthquake of  $M_W$  7.5. Geophys. Res. Lett. 29 (16).
- Tanaka, S., Sato, H., Matsumura, S., Ohtake, M., 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.