Location of river-induced seismic signal from noise correlation functions

A. Burtin, Jérôme Vergne, Luis Rivera, P. Dubernet

Geophysical Journal International

Geophys. J. Int. (2010) **182,** 1161–1173 doi: 10.1111/j.1365-246X.2010.04701.x

Location of river-induced seismic signal from noise correlation functions

A. Burtin,[1] J. Vergne,[2] L. Rivera[2] and P. Dubernet[1]

[1] *Laboratoire de Géologie, École Normale Supérieure de Paris, CNRS UMR* 8538, *Paris, France. E-mail: burtin@geologie.ens.fr*
[2] *Institut de Physique du Globe de Strasbourg (CNRS/UDS), Strasbourg, France*

Accepted 2010 June 14. Received 2010 June 11; in original form 2009 July 28

SUMMARY

We extend the use of noise correlation functions (NCFs) to locate stream segments of the trans-Himalayan Trisuli River that are responsible for the large high-frequency seismic noise observed at Hi-CLIMB stations. Cross-correlations of continuous seismic records at several pairs of stations indicate that some seismic sources are coherent only during the monsoon period. To locate these sources, we perform a migration of a selection of NCF envelopes filtered at frequencies ranging from 2 to 6 Hz. We produce a set of coherence maps at each frequency and for various apparent velocities to determine the regions which best explain the observed NCFs. The highest coherences are always located along restricted portions of the Trisuli River and are generally obtained for an apparent velocity of $3 \pm 0.3\,\mathrm{km\,s^{-1}}$. We also carry out a set of synthetic tests based on a full forward modelling approach for different distribution of sources. These simulations indicate that (1) the observed NCFs are dominated by Rayleigh surface waves and/or *S* waves, (2) cannot result from the coherence of *P* waves and (3) the recorded sources are effectively located along the Trisuli River. These tests also reveal some artefacts induced by the linear geometry of the Hi-CLIMB network. We take these artefacts into account and determine that the sources of seismic noise are mostly concentrated along the steepest portion of the Trisuli River with a maximum at the front of the High Range. We associate the river sources to the impacts of sediment particles on the channel bed and their distributions are in good agreement with incision rates along the river. Therefore, this study reveals the ability of locating zones of high river sediment transport and bedrock erosion based on the analysis of seismic noise recorded outside the stream.

Key words: Interferometry; Geomorphology; Tectonics and landscape evolution; Asia.

1 INTRODUCTION

The revolution brought by the digital acquisition of seismological data has increased the interest of studying continuous seismic recordings which has led in turn to an expansion in the exploration and comprehension of seismic noise. Ambient noise affects the entire frequency band of interest in seismology. At very low frequencies (0.002–0.01 Hz), the seismic 'hum' of the Earth has a complex origin which includes interactions between atmosphere, ocean and seafloor (e.g. Rhie & Romanowicz 2004, 2006). In the microseisms frequency band (0.04–0.5 Hz), seismic noise originates from swell activity in the deep ocean and along the coast (e.g. Longuet-Higgins 1950; Gutenberg 1958; Bromirski & Duennebier 2002; Kedar *et al.* 2008). At higher frequencies (>1 Hz), ambient noise is usually dominated by human activities such as road traffic and industries in the vicinity of stations (McNamara & Buland 2004). However, some local natural phenomena can drastically increase the high-frequency content of seismic noise. Among these factors, Burtin *et al.* (2008) have confirmed with the example of the Trisuli River in the Himalayas that turbulent streams are a major source of seismic noise at nearby stations.

From the spectral analysis of the continuous seismic signals recorded at a subset of stations from the Hi-CLIMB experiment, Burtin *et al.* (2008) observe a large increase of seismic energy in the 1–20 Hz frequency range during the summer monsoon period. The origin for this high level of high-frequency energy is attributed to the turbulence of the stream combined to the ground vibrations generated by the sediment transport. However, the individual spectral analysis cannot provide information about the location of the river portions where the seismic noise originates and which denote a higher transport of sediment and potentially a higher abrasion of the river bedrock. Sediment transport and bedrock incision are critical parameters in studies of the landscape evolution and for the mitigation of flood events. But their assessments with classical geomorphic methods are not always possible. Indeed, most of these approaches are *in situ* techniques and cannot be used during extreme hydrologic conditions. Thus, no information is available during flood events when most sediments are mobilized and

erosion processes happen. The spatial and temporal monitoring of river sediment transport with seismological sensors represents an alternative to *in situ* methods with no hydrologic dependency. In this study, we explore a first attempt to locate the sediment transport at a river scale based on the record of high-frequency seismic noise produced by the stream.

Location of sources that produce continuous high-frequency seismic noise is usually performed with small aperture arrays to retrieve the azimuth of the incoming waves (e.g. Rost & Thomas 2002). In the case of the Trisuli River, the geometry of the Hi-CLIMB network and the instrumentation are clearly not adapted to apply such array analyses. Here, we explore another technique, based on the use of noise correlation functions (NCFs), to locate the origin of this high-frequency seismic noise. NCFs rely on a theory developed by Weaver & Lobkis (2001) which indicates that noise records at a pair of receivers can be used to retrieve the empirical Green's function of the medium between these receivers. Over the last decade, this technique has been widely used to study the structure of the Earth lithosphere (e.g. Shapiro & Campillo 2004; Sabra *et al.* 2005; Yao *et al.* 2006). Most of these studies use the coherence of the seismic noise at low frequencies (∼0.01–0.1 Hz) because the correlations of long period waves are efficient for large interstation distances and more useful to probe the Earth interior. The validity of the theory strongly relies on a spatially homogeneous distribution of seismic noise sources or scattering points (e.g. Snieder 2004; Roux *et al.* 2005b). However, several observations have revealed an asymmetry in the NCFs, which indicates the existence of strong persistent and localized sources of seismic noise (e.g. Stehly *et al.* 2006; Pedersen *et al.* 2007; Yang & Ritzwoller 2008). NCFs have thus also been used to study the origin of the low-frequency ambient seismic noise (Shapiro *et al.* 2006; Rhie & Romanowicz 2006; Stehly *et al.* 2006; Brzak *et al.* 2009). The location methods include back projection of NCF maximum amplitudes (e.g. Stehly *et al.* 2006), grid search misfit (Shapiro *et al.* 2006) or migration of NCF envelope (Brzak *et al.* 2009). Here, we extend the latter technique to higher frequencies (>1 Hz) for locating the noise sources that produce the increase of seismic energy along the Trisuli River during the monsoon period.

After a brief description of the analysed seismological data, we present the procedure used to compute the NCFs and we detail their temporal fluctuations during the year 2003. With a selection of NCF envelopes, we proceed to the location of the seismic noise sources and we carefully analyse our results with a full forward modelling approach where we simulate various sets of source distributions. This allows highlighting the artefacts introduced by the geometry of the network and comparing our results from the distribution of river induced seismic noise with river incision rates.

2 HI-CLIMB DATA SET AND SEISMIC NOISE ANALYSIS

We analyse the continuous seismic signal at stations from the Hi-CLIMB experiment (Fig. 1), a temporary seismological network deployed across the India-Asia collision zone (Hetényi 2007; Nábělek *et al.* 2009). We use the same data set as described in Burtin *et al.* (2008), focusing our analyses on the data acquired during the year 2003 at stations H0330 to H0580, all installed at distances from the Trisuli River ranging from 0.1 to 2 km (Fig. 1). Stations H0410–H0500 are along narrow river segments with the steepest stream gradients and where the transport capacity is expected to be large. North of this segment (stations H0510–H0580), the slope

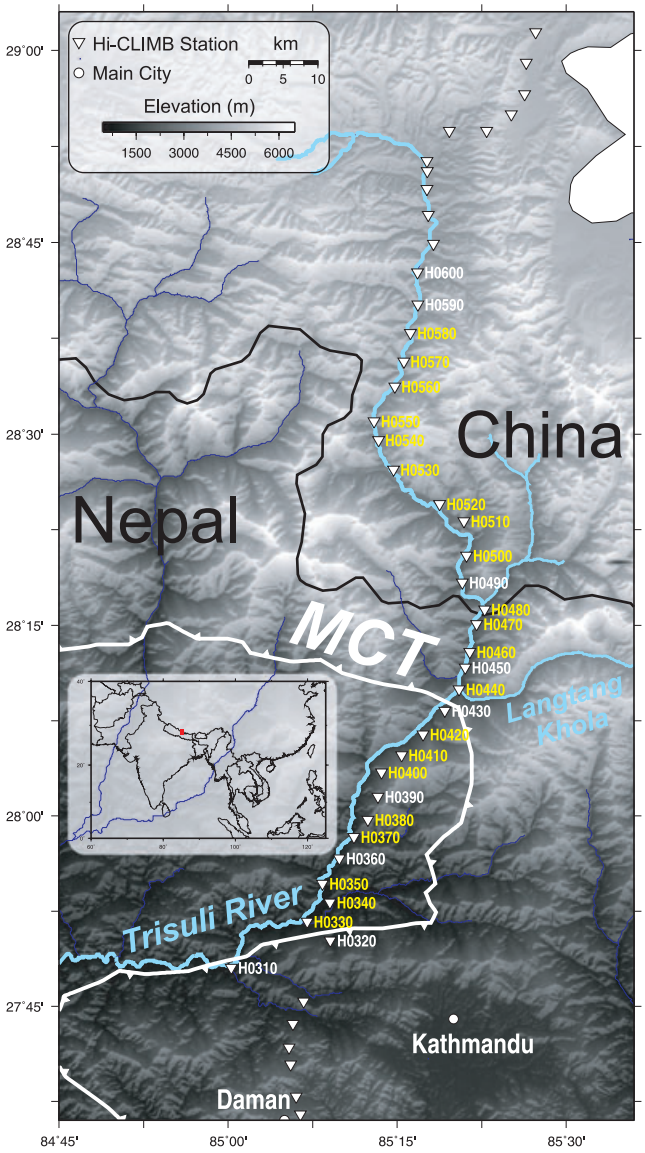

Figure 1. Location of the study area in the Himalayas of Nepal. Inverse white triangles represent the seismological stations from the Hi-CLIMB experiment. Station names are indicated in yellow for stations that are actually used in the Hi-CLIMB data set and in white when we include all possible sites in the synthetic data set. MCT refers to Main Central Thrust.

of the river is lower and the supply of water is reduced due to a snow and rain shadow of the High Range (Putkonen 2004). To the south (stations H033–H0400), the stream slope is also weak and the river becomes wider with meanders. All Hi-CLIMB stations were equipped with broad-band seismometers and 24-bit recording systems continuously operating at a sampling frequency of 50 Hz with an antialias low-pass filter at 20 Hz. Despite some instrumental failures, most of the analysed seismological stations provide long enough records during the year 2003 to extract sufficient coherence between stations.

2.1 Determination of noise correlation functions

The computation of the NCFs for each pair of receivers involves several pre-processing steps described in detail by Bensen *et al.* (2007). We basically follow the same procedure that can be

summarized by: (1) For each station, we split the continuous records of the vertical component into daily long signals. (2) We deconvolve each signal from the instrument response and remove the mean and the trend. (3) We apply a bandpass filter in the 1–20 Hz frequency band that roughly corresponds to the band for which an increase of seismic noise is observed during the summer monsoon (Burtin *et al.* 2008). (4) We apply a temporal normalization to reduce the perturbations induced by earthquakes, instrumental issues or any punctual sources of seismic noise. Among the various types of normalization, we choose the 1-bit normalization which consists of keeping only the sign of the time series. (5) For each station pair, and for each day when the two stations are operating simultaneously, we compute the cross-correlation of the two time-series (hereafter the NCF) and for time delays ranging from −40 to +40 s. (6) Finally, because the NCFs have a broad frequency content (1–20 Hz), we compute several filtered NCFs by applying a series of Gaussian filters centred around frequencies ranging from 2.75 to 10 Hz and with a standard deviation σ_f of 0.25 Hz. These filtered NCFs will be used in the migration process to locate sources of seismic noise which produce coherent energy at each frequency.

2.2 Time evolution of noise correlation functions

We present in Fig. 2(a) an illustration of a 1-day raw NCF for the station pair H0420–H0460 (∼14 km apart). We observe the emergence of a signal at positive lags between 3 and 8 s but with a poor signal-to-noise ratio (SNR). Stacking all the 1-day NCFs, for 2 months in the summer season (July and August), clearly increases the SNR and confirms the presence of pervasive and powerful sources of seismic noise (Fig. 2b). Furthermore, the asymmetry of the NCF between the acausal and causal part (negative and positive lags) indicates an inhomogeneous source distribution with a preferential azimuth for the incoming seismic waves. The stacked NCF computed during the off-monsoon period is clearly different from the monsoon period one. This illustrates that the sources of seismic noise are not coming from the same region over the year and that they are less coherent and powerful during the off-monsoon period.

To assess the temporal variability of the coherence, we look at daily NCF envelopes over the time for various station pairs and at several frequencies. For the same couple H0420–H0460 the NCF filtered around 3.5 Hz shows that the two peaks between positive lags of 3 and 8 s appear around mid-June (Fig. 3a). They can be observed almost every day of July and August and perhaps until September, although some instruments failures preclude a continuous sampling in this time period. Later during the year, both peaks are no longer visible. It indicates that the coherent signal is generated only during the monsoon period when the discharge of the Trisuli River is the highest. The presence of these peaks also indicates that there are two major zones of seismic noise generation with a different azimuth.

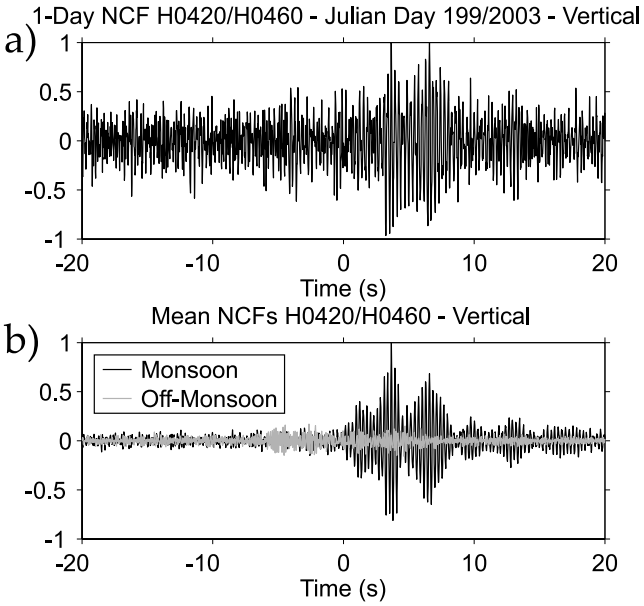

Figure 2. (a) Raw vertical daily noise correlation function between H0420 and H0460 (Fig. 1) for the Julian day 199. Some coherence begins to appear for positive delays between 3 and 8 s. (b) Mean raw noise correlation functions during the monsoon (black) and the dry season (grey) for the same station pair. Their amplitudes are normalized to the maximum of both curves. We clearly observe strong peaks of correlation whereas during the dry season no coherence is detected.

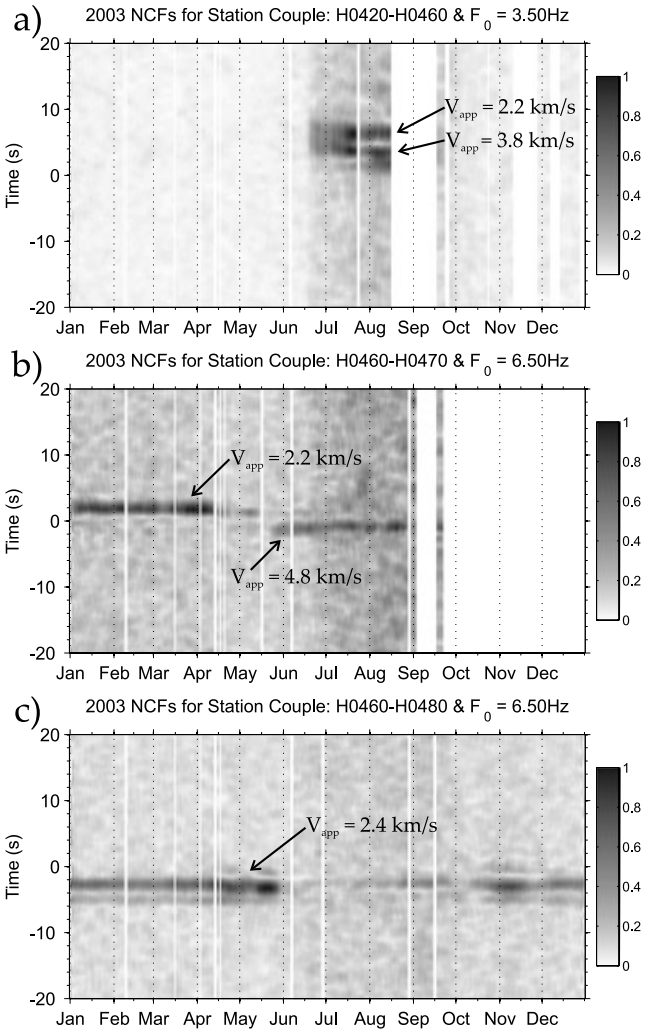

Figure 3. Yearly evolutions of the noise correlation function envelopes for three station couples: H0420–H0460, H0460–H0470 and H0460–H0480 for (a), (b) and (c), respectively. For each station pair, the amplitudes are normalized to the yearly maximum. In (a) and (b), the summer peaks of correlation are associated to the river seismic noise, whereas in (b) and (c) the off-monsoon coherence is associated to anthropogenic sources.

Indeed, if sources were located along the median between both stations, they would produce a peak of amplitude at 0 s lag on the NCF, corresponding to an infinite apparent velocity. In opposite, the maximum absolute lag would be noticed if sources were aligned with the stations. Thus, relative to the station pair direction, the coherence with an apparent velocity of 2.2 km s^{-1} is coming from a smaller azimuth than the one with an apparent velocity of 3.8 km s^{-1} (Fig. 3a).

Another typical example of the NCF variability is given by the station pair H0460–H0470 (∼4 km apart), where the filtered envelopes at 6.5 Hz denote good coherence during the entire year 2003. However, we observe that prior to the beginning of the monsoon the maximum of correlation occurs at a positive lag with an apparent velocity of 2.2 km s^{-1} whereas during the monsoon the maximum of correlation is for a negative lag and an apparent velocity of 4.8 km s^{-1} (Fig. 3b). This observation reveals a different origin for the sources of seismic noise over the year. We interpret the pre-monsoon peak of correlation which disappears during the summer to be associated with anthropogenic sources. Indeed, human activities such as industries or road traffics are dominant in the off-monsoon period. During the rainy season, the large rainfalls provoke a drastic decrease of anthropogenic activities. In opposite, the peak that emerges in summer has to be considered as an indication of river induced seismic noise. Finally, Fig. 3(c) shows the yearly fluctuation of the filtered NCF envelopes at 6.5 Hz for the stations pair H0460–H0480 (∼6.5 km apart). This is an example of seismic noise sources that are more energetic during the off-monsoon period. Such behaviour is probably representative of anthropogenic sources.

2.3 Selection of noise correlation functions

Our interest is in the location of the sources which produce the burst of seismic noise during the monsoon period, and especially the segments of the Trisuli River that are the most active. To discard the non-desired sources as the anthropogenic activities, some basic tests on NCFs are necessary to discriminate between river and human made noise. Thus, several selection processes have to be applied on the NCFs, to insure that we will only use NCFs for which this particular noise is observed. First, we exclude station pairs that have an interstation distance larger than 30 km. According to observations, we do not notice any coherence at such large interstation distances and we suspect that the attenuation of the high-frequency signal is too important for large station couple distances. In addition to this distance threshold, we use a temporal threshold that excludes stations couples that have less than 25 days in common to average. Afterward, we seek for the existence of a relevant peak of correlation with the help of an SNR criterion. The SNR amplitude of an NCF envelope is computed from the ratio of its maximum amplitude relative to its standard deviation. We explored different values and an SNR threshold of 3.5 appears as a good compromise to exclude non-relevant data and keep a significant amount of station pairs. Finally, we impose a seasonal condition to look for peaks of coherence that show an increase of their amplitude during the monsoon period. For each station pair and each frequency, we compute the seasonal fluctuation $\Delta H(t) = H_{\mathrm{m}}(t) - H_{\mathrm{off-m}}(t)$, where $H_{\mathrm{m}}(t)$ is the NCF envelope corresponding to the stack of the daily NCF for the months of July and August, and $H_{\mathrm{off-m}}(t)$ the off-monsoon NCF envelopes for days out of a period going from May to October. We finally keep all the station couples for which the mean of ΔH is positive.

3 SOURCE LOCATION OF SEISMIC NOISE

3.1 Migration of NCF envelopes

For a set of NCF envelopes at a given frequency, we call 'migration', the process that through an exploration of the space around the network reveals the points that best explain the arrival time of the most prominent peaks. This approach is very similar to the graphical approach formerly used to locate earthquakes when only arrival times of P waves at a set of stations are available (Mohorovičić 1916). First for a pair of receivers, $H_{\mathrm{m}}(t)$ is normalized to its maximum and we consider a given point in a discrete space of constant velocity (Fig. 4). For a source located at this point, we calculate the theoretical time-delay dt_{calc} between the arrival time of the wavefront at both stations and we associate to this hypothetic source point the amplitude we observe at this time-lag given by $H_{\mathrm{m}}(dt_{\mathrm{calc}})$. This operation is performed for each point of the discrete space and, because the wave propagation velocity is unknown, we repeat this for a set of apparent velocities ranging from 1 to 5 km s^{-1}. For a given velocity, one time-delay dt_{calc} will correspond to a possible set of sources that are located along a hyperbola for which the two stations are the focal points. Only NCF amplitudes, corresponding to $dt_{\mathrm{calc}} \in [-D/V, +D/V]$, where D is the distance between the two receivers and V is the apparent velocity, will be migrated because $|D/V|$ is the observed maximum time-lag for sources that are aligned with the stations. Fig. 4(c) illustrates the migration of a NCF envelope for the couple H0420–H0460, filtered around 3.5 Hz, and assuming an apparent velocity of 3 km s^{-1}. The maximum of coherence on the NCF is observed at positive lags and is mapped into hyperbolic forms on a coherence map. The remaining ambiguity about the source locations will be removed when we migrate all the selected NCF envelopes. Therefore, we average all the coherence maps associated to each station couple and the best source location and apparent velocity are retrieved when we look for the maximum of coherence.

3.2 Source locations for Hi-CLIMB Data

Results of the migration for NCF envelopes at four different frequencies (3.5, 4, 4.5 and 6 Hz) are presented in Fig. 5. For a central frequency at 4 Hz (Figs 5a–c), we show the coherence maps for three velocities (1.5, 2.9 and 5 km s^{-1}) to illustrate the evolution of the coherence with the propagation velocity. For low velocities, the addition of the hyperbolas, which describe the maxima of NCFs envelopes, maps into a wide region of weak coherence located on each side of the stations. When the apparent velocity increases, the maximum coherence increases, and its area narrows and converges on segments located on the river. The maximum of coherence corresponds to an apparent velocity of 2.9 km s^{-1} (Fig. 5b) and two main segments of river are highlighted. One is located between stations H0400 and H0420 and the second between stations H0460 and H0480 (see Fig. 1 for station locations). When the velocity exceeds 3 km s^{-1}, the areas of maximum coherence tend to remain along the river but their amplitudes decrease (Fig. 5c). Figs 5(d)–(f) show the coherence maps for three other frequencies (3.5, 4.5 and 6 Hz) and for a velocity at which the maximum of coherence occurs. The sources of seismic noise are always located in the Trisuli River along a portion defined by stations H0400 and H0480. However, the imaged segments are slightly different with the frequency which may indicate various mechanisms of noise generation along the river. For example, the patch located between stations H0460 and H0480

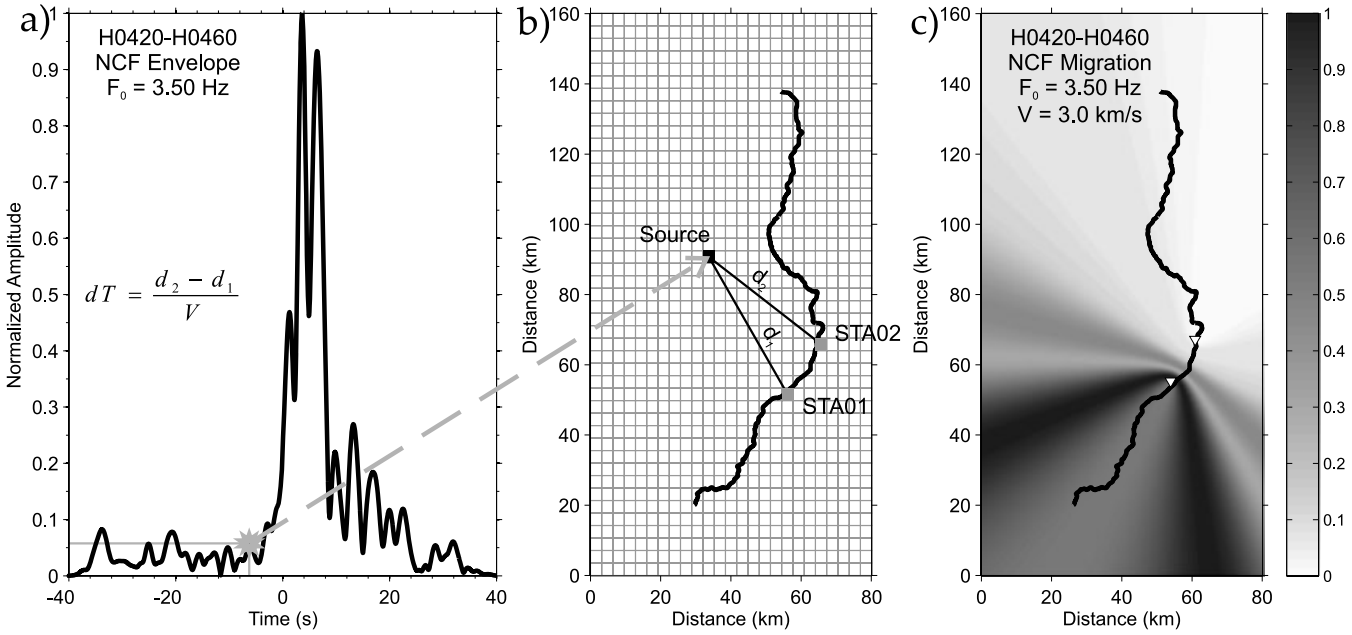

Figure 4. The source migration process. (a) Monsoon envelope of the Noise Correlation Function between H0420 and H0460 at 3.5 Hz. The amplitudes are normalized to the maximum. (b) Discrete space where each grid point will be assimilated as a potential source of noise. For a given point, velocity and station couple, we compute the theoretical delay and we associate to this given point the envelope amplitude observed at the calculated delay. (c) Coherence map for the migration of the H0420–H0460 envelope at 3.5 Hz. The inverse white triangles represent the two seismological stations.

is only observed for frequencies higher than 4 Hz whereas the one between H0400 and H0420 is seen for all frequencies. In addition, many coherence maps show a systematic northward spreading of large coherences along the main orientation of the Hi-CLIMB profile (Fig. 5). This constant pattern may indicate a bias that is introduced by the linear geometry of the array. Nevertheless, the sources of seismic noise in summer seem well located along the Trisuli River.

We perform a systematic analysis of the migrated maps obtained for various frequencies ranging from 2.75 to 10 Hz and propagation velocity ranging from 1 to 5 km s^{-1}. The number of selected NCF envelopes with frequencies strongly decreases for frequencies higher than 5 Hz (Fig. 6a). This decrease is associated with an increase in the variability of the best velocity, referred as the apparent velocity corresponding to the overall maximum coherence. For frequencies lower than 5 Hz, best velocities remain almost constant around 3 km s^{-1} (Fig. 6b). At higher frequencies, the fluctuation of the velocity is more random and may reveal a lack of correlation in seismic signals and a weak ability to locate noise sources. If we consider only frequencies for which at least 15 station pairs are used, velocities as well as noise locations are coherent. We see however some exceptions for frequencies of 5 and 6.25 Hz that both give a best solution for a much larger velocity (>5 km s^{-1}). The evolution of the maximum of coherence with the apparent velocity indicates that velocities around 3 km s^{-1} tend to produce an almost equivalent peak of correlation amplitude compared to velocities around 5 km s^{-1} (Fig. 6c). Apart from these last two frequencies, the reliable coherence maps give a best velocity of 3 ± 0.3 km s^{-1}. This relatively low value cannot correspond to the velocity of *P* wave because no thick sediment layer exists in the region of focus. Based on the crustal velocity structure published for the Himalayas (Monsalve *et al.* 2006), our best propagation velocity rather fits with Rayleigh surface wave and *S*-wave velocities of the upper layer (Fig. 7).

4 MODELLING THE SOURCES OF SEISMIC NOISE

To test the robustness of the location method, identify the wave content that makes the NCFs and justify the use of a constant migration velocity, we carry out a full forward modelling approach by generating synthetic NCFs for various distribution of sources. Several studies about ambient noise correlation have already tried to reproduce synthetic NCFs by stacking a multitude of synthetic seismograms computed for a generally stochastic distribution of sources in space and time (e.g. Yang & Ritzwoller 2008; Brzak *et al.* 2009). In many cases, the synthetic seismogram corresponding to a single source is a wavelet shifted in time according to a constant propagation velocity. If such an assumption seems realistic for low-frequency wave contents and large source-station distances, it might not be appropriate for high-frequencies and near field studies. In addition, this assumption does not allow exploring the full wave content of seismic noise.

4.1 Simulation parameters

To simulate the NCFs between a pair of stations, we compute one synthetic time-series of ambient seismic noise at each station, by stacking seismograms produced by a distribution of punctual seismic sources. The synthetic seismograms are obtained from a full waveform modelling approach (Herrmann 2002). Based on a given 1-D velocity structure and source-station distance, we calculate the corresponding Green's function which is then convolved with a source time function. To reduce the computing-time, we only calculate the theoretical Green's functions for distances ranging from 0.05 to 100 km with 0.05 km spacing. This value is small enough compared to the smallest wavelength than can be recorded by the Hi-CLIMB data set. For the computation of Green's functions, we select a velocity model that is representative of the Himalayas of

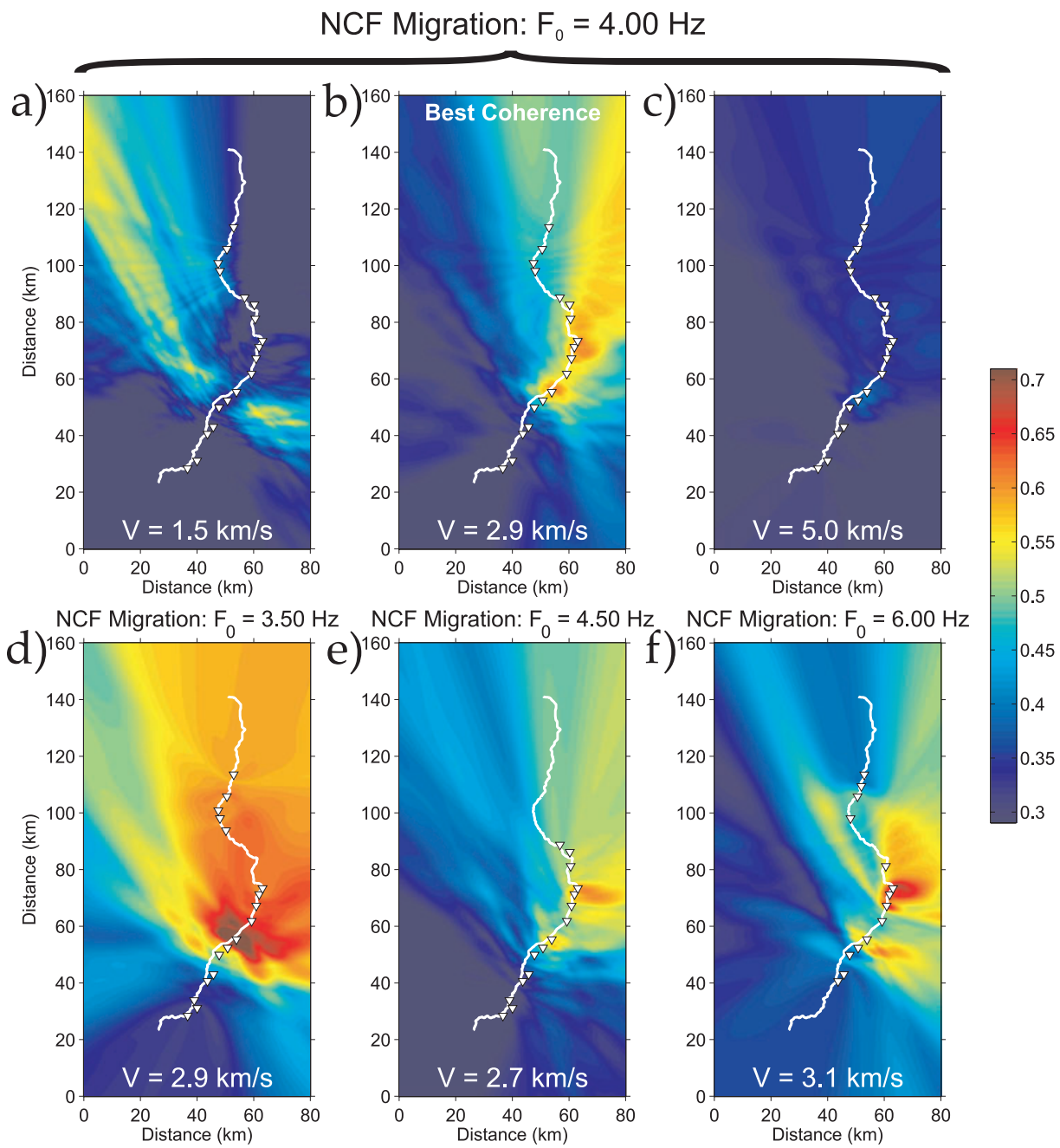

Figure 5. (a–c) Coherence maps at 4 Hz and for an apparent velocity of 1.5, 2.9 and 5 km s^{-1} from left to right, respectively. The explored apparent velocity is indicated at the bottom of the map. The best source locations at 4 Hz are associated with the velocity of 2.9 km s^{-1}. (d–f) Best coherence maps for a frequency of 3.5, 4.5 and 6 Hz from left to right, respectively. Best found velocities are indicated at the bottom of the map. Mean correlation amplitudes are shown in each coherence map. Red colours stand for areas of large coherence (most probable source of noise), whereas blue colours stand for areas of low coherence. The inverse white triangles represent the seismological stations actually used in the migration. Note the systematic northward spreading of the high coherence zones.

Nepal (Monsalve *et al.* 2006) and illustrated on Fig. 7. The velocities of *P* and *S* wave in the upper crust are 5.6 and 3.2 km s^{-1} and we do not add any low velocity layer on top. For each simulation, we generate 6 h of seismic records with a statistical rate of 10 sources per second (random time-distribution). The amplitude modulation of the vertical impulsive forces spans three orders of magnitude. We then compute the simulated NCFs following the same procedure as for the treatment of the Hi-CLIMB data set (Section 2.1).

The first simulation (named hereafter S1) consists of 70 per cent of seismic sources uniformly distributed along a single river segment and 30 per cent homogeneously distributed in the studied area presented in Fig. 8. The aim of this simulation is to test the ability of retrieving sources that are distributed along the Trisuli River. In the second simulation (S2), 50 per cent of the sources originate from a point location (Fig. 8). In this manner, we want to estimate the potential of locating a persistent source point. Such a test can illustrate

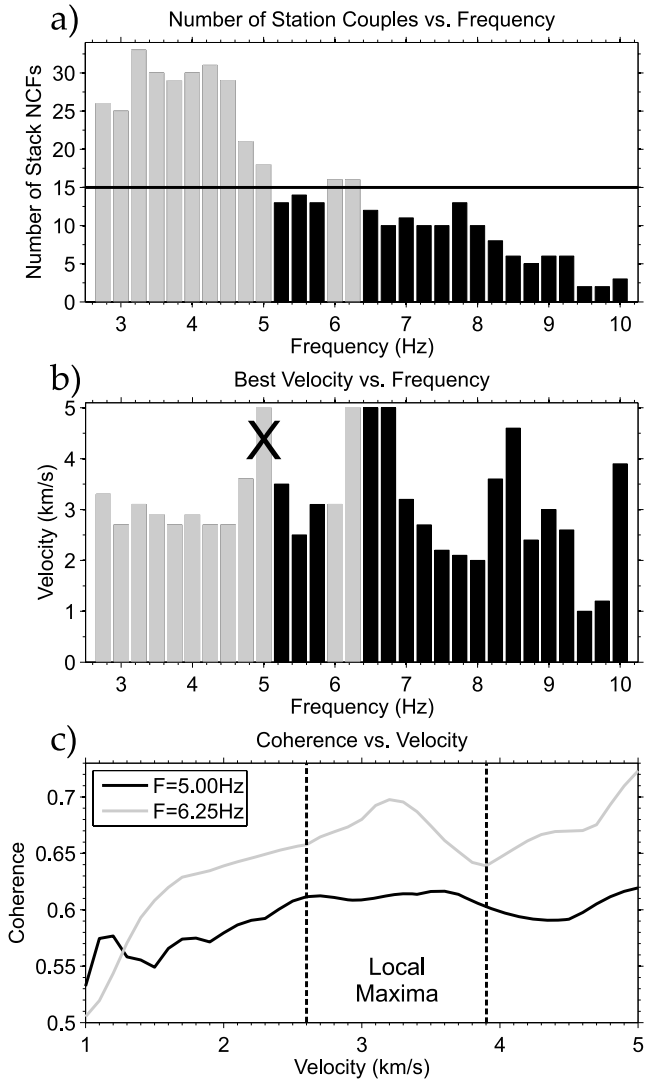

Figure 6. (a) Number of station couples used in the source migration process in frequency. The horizontal black line marks a number of 15 station pairs at least used in the location process (grey bars). For frequencies higher than 4.5 Hz, the number of selected NCF envelopes decreases rapidly. (b) Best apparent velocity found from the migration of NCFs in frequency. Grey bars stand for the frequencies that offer confident results. With the exception of two of these frequencies (5 and 6.25 Hz), we find a mean velocity of about 3 km s^{-1}. The 5 Hz frequency is discarded in the discussion section due to a poor resolution of the best apparent velocity. (c) Maximum of the coherence map as a function of the apparent velocity of the medium at 5 and 6.25 Hz. These two curves show the local maxima that we can observe between 2.6 and 3.9 km s^{-1} (vertical dash lines) and for which the best velocity agrees with the other frequencies.

the effect of anthropogenic perturbations which are usually much localized. The simulation S3 is performed to reproduce the ambient seismic noise of two separate uniform river segments, where 30 per cent of the total sources are randomly dispersed (Fig. 8). Finally, to better constrain a geometrical effect of the Hi-CLIMB array, we make a simulation (S4) where all sources are randomly distributed in the plane (Fig. 8). During the location process, we realize two kinds of migrations: one that uses only the station couples that are actually selected from the Hi-CLIMB data at 3.5 Hz and a second that uses all the possible station pairs that have an interstation distance smaller than 30 km.

4.2 Simulation results

We illustrate the results of the four simulations by showing the migration of the NCF envelopes filtered at 3.5 Hz. Similar observations and conclusions can also be made for other frequencies. Fig. 9 presents the migration for the simulation S1 (sources located along a segment of the river stream) for different apparent velocities. The overall pictures share several similarities with the ones deduced from real data (Figs 5a–c). When the velocity is underestimated, the coherence is low and coherent regions appear on each side of the profile. The maximum of coherence is reached for an apparent velocity of 3 km s^{-1} (Fig. 9b), a value similar to the one obtained with Hi-CLIMB data and slightly lower than the *S*-wave velocity in the upper part of the model. Moreover, for this particular velocity, the area of maximum coherence mostly corresponds to the river segment where we impose the seismic sources. We also denote a northward spreading of the area of strong coherences. This spreading tends to be reduced when we use all the available stations couples (Fig. 9d) which indicate a possible artefact caused by the linearity of the network and the lack of retained NCFs from H0490 to H0520. When the apparent velocity exceed 3 km s^{-1}, the overall coherence diminishes but the coherent regions remain on the same location (Fig. 9c).

Fig. 10 presents the results of the simulations S2, S3 and S4 for the apparent velocity corresponding to the maximum of the overall coherence that is always between 2.8 and 3.2 km s^{-1}. In the case of the simulation S2, the point source that we impose is well retrieved whether we consider all the couples or a subset of them (Figs 10a–b). The area that delineates the best solution is nevertheless better constrained with a larger number of station pairs. The 50 per cent randomly distributed sources have no effect on the result of the migration which demonstrates that a localized persistent source of noise can be recovered in spite of being a non-unique source. In the simulation S3 and for the migration of a full set of stations, we clearly retrieve the two river segments (Fig. 10d). With a reduced set of stations, the migration of NCFs mostly reveals the southernmost segment of river (Fig. 10c). This segment is actually shifted northward close to H0420 compared to the imposed geometry of sources. Moreover, unexpected coherence is observed between the two segments. This test points the importance of the station location relative to the sources. Close stations are helpful to delineate the spatial distribution of seismic noise sources, especially with the Hi-CLIMB linear geometry. In the simulation S4, the migration of a random distribution of sources infers that the area of maximum coherence roughly mimics the shape of the Trisuli River. When we use a reduced set of stations, this area extends along the main direction of the array, but outside the network (Fig. 10e). Using all station couples (<30 km), the image exhibits almost similar results with the exception of a stronger coherence on the Trisuli River between stations H0410 and H0510 (Fig. 10f). The recovered distribution of sources is clearly an artefact of the linear geometry of the seismic array. Indeed for a single station pair, only the sources aligned along the direction of the couple tend to produce seismic signals which stack coherently in the NCF. Thus, the migration of such a NCF will indicate a zone of strong coherence on each side of the segment that links the two stations. For the Hi-CLIMB network, since all station couples have an almost equivalent azimuth, the area of strong coherency follows the same direction. For the real data,

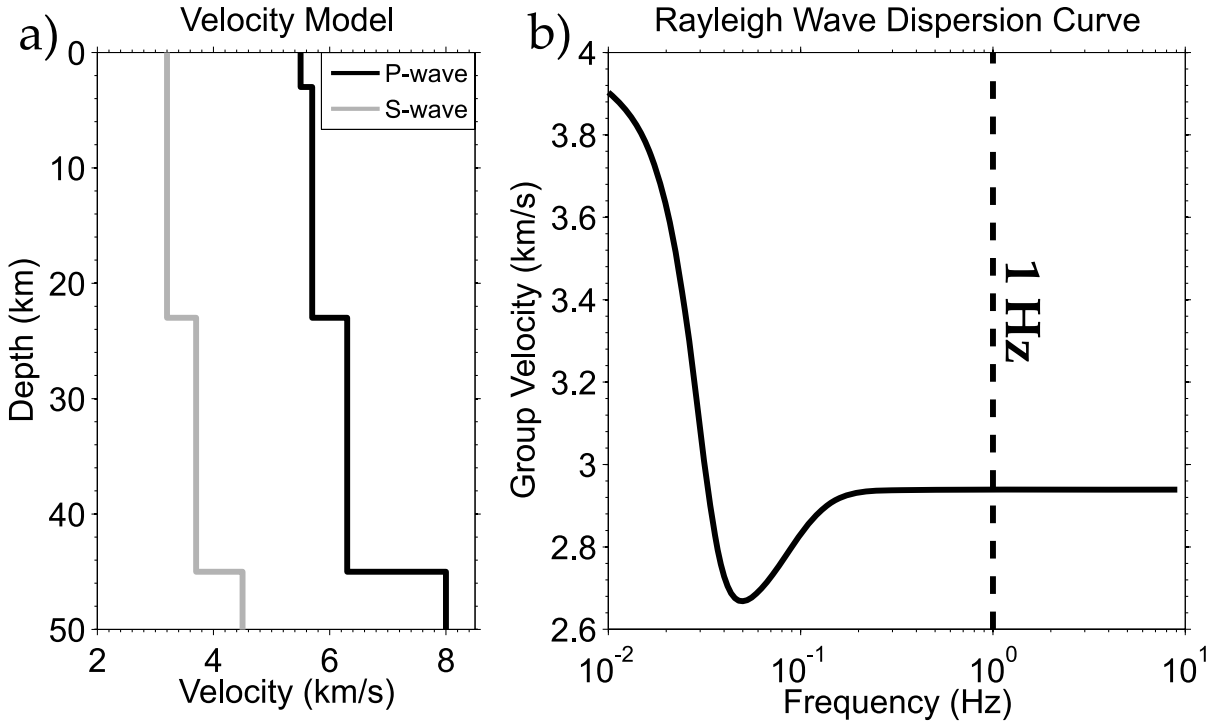

Figure 7. (a) Velocity structure of the Himalayas of Nepal determined by Monsalve *et al.* (2006). *P*- and *S*-wave velocities are illustrated by the black and grey lines, respectively. (b) Rayleigh wave dispersion curve for the given model in (a). Note that for the frequencies of interest (>1 Hz), the group velocity of surface wave is constant with a value of 2.9–3 km s^{-1}.

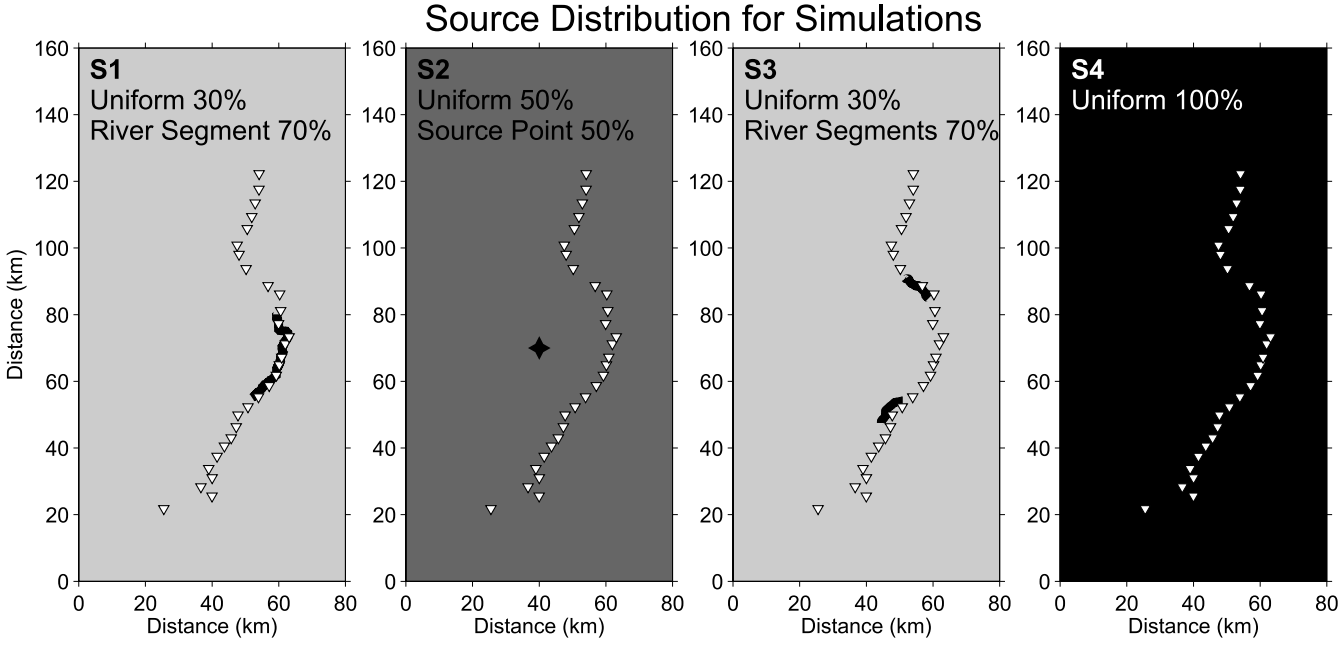

Figure 8. Source distribution for the simulated seismic noise S1–S4 from left to right, respectively. For the simulation S1, a single river segment (black) produces uniformly distributed river sources. For the simulation S2, the black cross marks the location of a source point. For the simulation S3, two river segments (black) produce uniformly distributed river sources. For the simulation S4, all seismic sources are uniformly dispersed in the medium. Inverse white triangles represent Hi-CLIMB stations.

we do not observe any region of strong coherences in the southern prolongation of the array and for some frequencies in the northern prolongation of the network (Fig. 5). Compared to the simulation S4, the coherence we notice on the data do not seem to be generated by a random distribution of sources. However, to fully rely in our observations we must take the geometric artefact into account.

4.3 Propagation velocity and wave content

Results from the treatment of real data and from all the simulations indicate that the apparent velocity associated with the overall best coherence is quite constant with frequency and ranges between 2.8 and 3 km s^{-1}. The theoretical dispersion curve of the regional

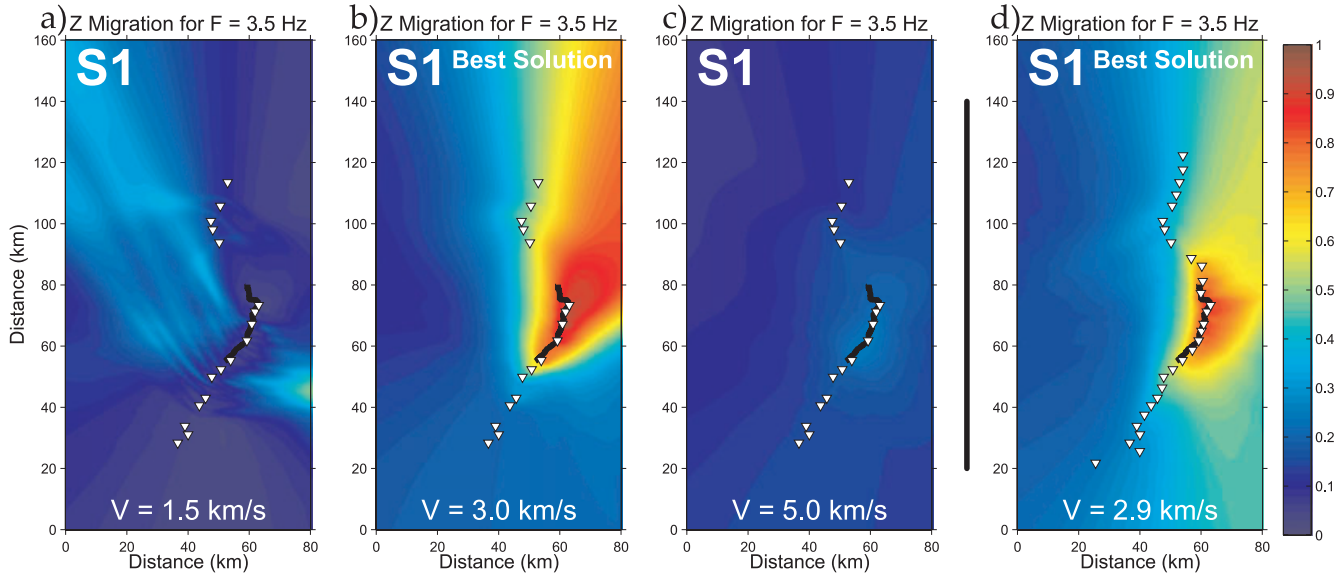

Figure 9. (a–c) S1 migrations at 3.5 Hz for a subset of stations at a velocity of 1.5, 3 and 5 km s^{-1}, respectively. The best solution is given for an apparent velocity of 3 km s^{-1} (b). (d) The best solution for S1 migration at 3.5 Hz for a full set of stations at a velocity of 2.9 km s^{-1}. The mean correlation amplitude is shown in each coherence map. Red colours stand for areas of strong coherence (most probable source of noise), whereas blue colours stand for areas of low coherence.

velocity model from Monsalve *et al.* (2006) shows for frequencies greater than 1 Hz a constant group velocity of 3 km s^{-1} (Fig. 7b). We also do not have any evidence for a focusing of coherence at higher velocities (tests performed up to 8 km s^{-1}) that would correspond to the correlation of *P* waves in the NCFs. Nevertheless, Roux *et al.* (2005a) shown that *P* waves can produce peaks of correlation when using a small aperture array of seismometers. To understand which type of seismic wave produces the coherence in the NCFs, we use the synthetics from the simulation S4 to compute the Green's function between two stations with an interstation distance of 15.5 km. In Fig. 11, we superpose on this empirical Green's function, the theoretical Green's function corresponding to an equivalent source-station distance. According to the used velocity structure, we do not detect a *P* or *S* wave on the empirical Green's function. The *S* wave tends actually to be overwhelmed by the amplitude of the Rayleigh surface wave on the theoretical Green's function. We carried out similar comparisons for a distance range of interest in our study and they all indicate that the Rayleigh surface wave is dominant on the NCFs. However in our simulations, the 1-D regional velocity model is not appropriate to describe the full complexity of the wave propagation. At high frequencies, the wave content of the Hi-CLIMB data may reflect the influence of scattering points and a complex topography that are not taken into account during the computation of the synthetic seismograms. One could suspect some strong Rayleigh to *S*-wave conversions, and since their propagation velocities are relatively close, it may prevent us to discriminate between both types of waves. Thus, it is clear that the main peaks on the NCFs do not reflect a *P* wave content but rather correspond to *S* waves and Rayleigh surface waves. Despite this remaining ambiguity, all these observations explain why we observe a relative constant apparent velocity in the investigated frequency range.

5 DISCUSSION

To highlight the segments of the Trisuli River that concentrate the sources of seismic noise, we extract the mean coherence along the stream from the best coherence maps at the reliable set of frequencies (Fig. 6). In the southern section of the stream, from H0330 to H0380, the coherence is the lowest which denotes very little dominant seismic sources (Fig. 12a). Farther north, from H0400 to H0510, the coherence is high with two main spots located around stations H0410–H0420 and H0470–H0480. North of H0520, the coherence diminishes but remains larger than south of station H0400. To partly remove the bias due to the geometry of Hi-CLIMB network, we compute the mean best coherence map that results from a homogeneous source distribution (simulation S4) and for the selected sets of stations and frequencies (Fig. 13). We extract the resulting coherence along the Trisuli River to obtain the geometric artefact of the array. We then divide the observed mean coherence along the stream (Fig. 12a) by the one obtained with the simulation S4. After applying this correction, the coherence remains the lowest south of H0380 and confirms the very few sources detected there (Fig. 12b). Northward, the high coherent zone near H0410 is preserved whereas the spot, initially observed around H0480, is not conserved anymore after the network correction. Quite large amplitudes are observed in the northern section of the river, between H0530 and H0550 and even enhanced after applying the network correction (Fig. 12b). We suppose this to be a residual artefact due to the weak number of selected NCF envelopes in the northern section, from H0520 to H0580 (Fig. 13). As inferred from the simulation S3 (Fig. 10c), the relative position of stations to sources is a relevant factor. Because the northernmost stations are not often selected in the location process, we have a poor resolution along this part of the river. In addition, a large amount of sources close to stations H0530 to H0550 during the monsoon period would produce a large increase in the seismic noise, a feature not observed on the spectrograms at these stations (Burtin *et al.* 2008). It may also explain why few NCFs are selected since little coherence is noticed there. In the following discussion, we discard the results obtained in the northernmost part of the river. However, these observations denote the ambiguities we have to deal when locating seismic noise sources with a linear profile of stations.

The active central segment (H0400–H0510) that concentrates most of the sources of seismic noise in the river is characterized by

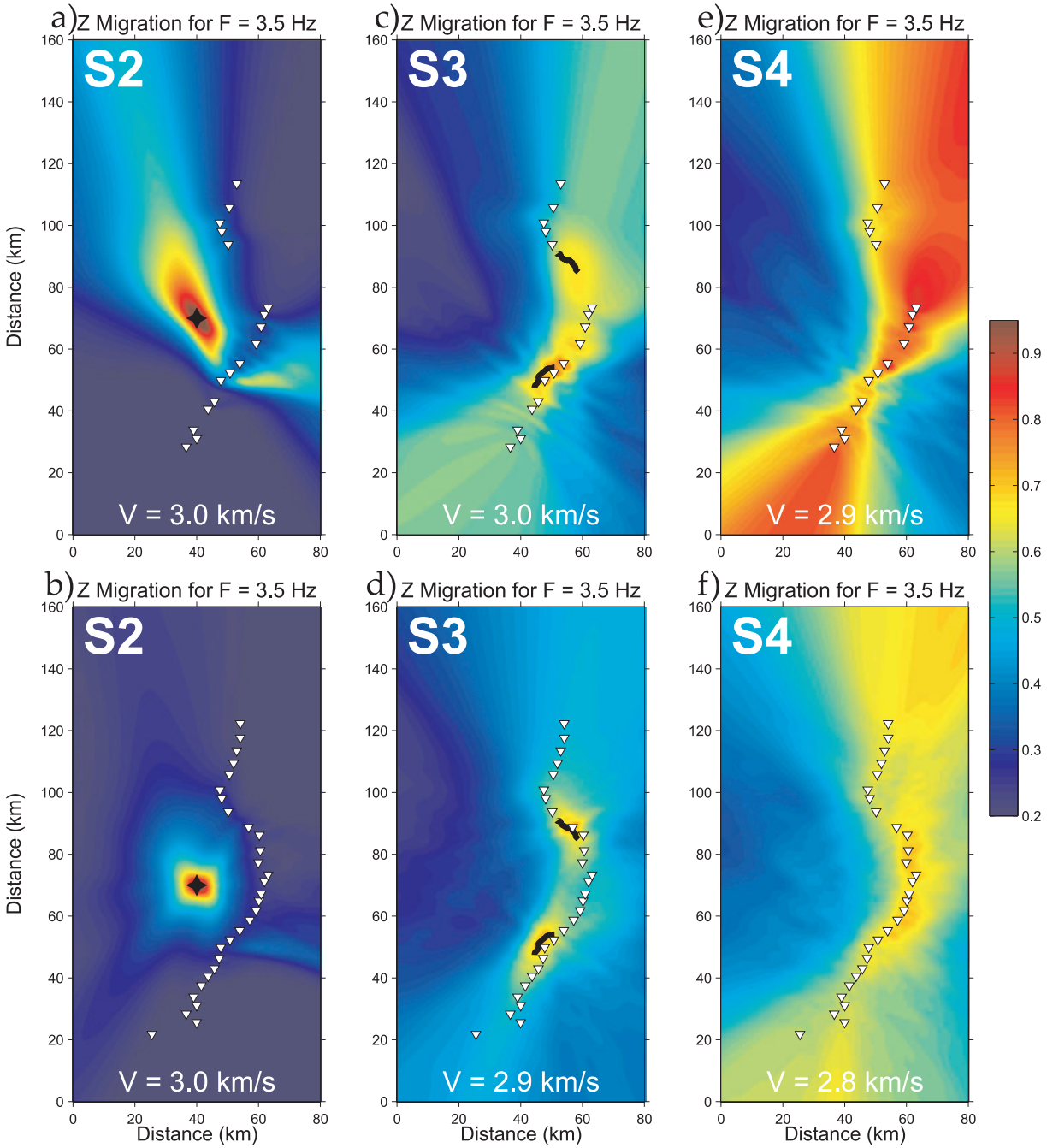

Figure 10. (a,b) Best solutions for S2 migration at 3.5 Hz for a subset and a full set of stations at both a velocity of 3 km s^{-1}, respectively. (c,d) Best solutions for S3 migration at 3.5 Hz for a subset of stations at a velocity of 3 km s^{-1} and for a full set of stations at a velocity of 2.9 km s^{-1}, respectively. (e,f) Best solutions for S4 migration at 3.5 Hz for a subset of stations at a velocity of 2.9 km s^{-1} and for a full set of stations at a velocity of 2.8 km s^{-1}, respectively. S2, S3 and S4 stand for a point source of noise (black cross), two small uniform river segments (black lines) and a homogeneous source distribution, respectively. The mean correlation amplitude is shown in each coherence map. Red colours stand for areas of strong coherence (most probable source of noise), whereas blue colours stand for areas of low coherence. Inverse white triangles represent the processed Hi-CLIMB stations.

a steep stream slope (Fig. 12). Burtin *et al.* (2008) have demonstrated that the variation of seismic noise was time-correlated with the variation of the water discharge. In addition, they inferred a hysteresis relation between the water level and the seismic energy pointing a link with sediment transport. A similar relation between suspended sediment loads and water discharges has also been observed along another trans-Himalayan river (Gabet *et al.* 2008). These observations indicate that the turbulence of the water is not the unique source of seismic noise and that sediment transport is partly responsible for the increase in seismic noise during the monsoon. Indeed, impacts of sediment particles on the channel bed produce ground vibrations that are recorded at nearby stations, increasing the observed seismic noise. Such seismic sources will mostly concentrate in region of steep stream slopes and large supply of water, where the hydrodynamic of the river is efficient to mobilize a large amount of sediment.

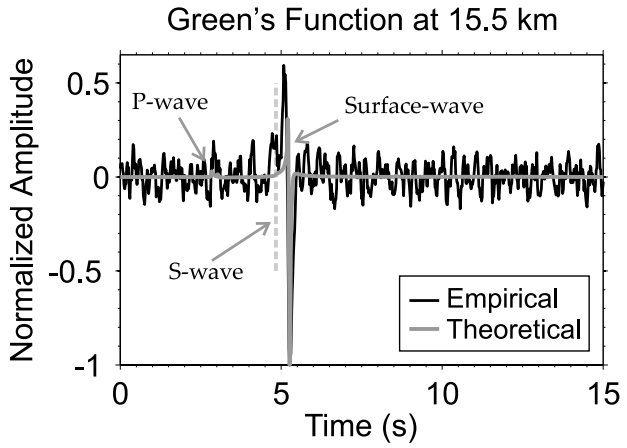

Figure 11. Comparison of the empirical (black) and the theoretical (grey) Green's function at 15.5 km. The empirical Green's function is estimated using the simulation S4 (homogeneous source distribution). The maximum of coherence is associated with the arrival of the Rayleigh surface wave.

The mechanical erosion caused by the impact of the bed load on the bedrock, called abrasion, is one of the major processes that control the landscape of a river and thus the relief. Lavé & Avouac (2001) have estimated the incision rates of several trans-Himalayan rivers along the Himalayan Arc from the long-term incision rates recorded by fluvial terraces. However, because no terrace is available along the steep gorges of trans-Himalayan rivers, they completed the erosion rate from a semi-empirical relation based on the estimated fluvial shear stress exerted by the flowing water on the stream bed (e.g. Howard *et al.* 1994; Whipple & Tucker 1999). In the case of the Trisuli, they observe that a peak of river incision (Fig. 12c) is located at the front of the High Range, where stations H0410 and H0420 were installed (Fig. 1). This well incised portion of river agrees with a segment of the largest coherence, hence the largest concentration of seismic sources. This region is also the locus of many debris flows (Burtin *et al.* 2009) that concentrate in the frontal part of the High Range due to the presence of loose soil and important rainfall (Thouret 1983; Marston *et al.* 1998). These hillslope processes participate to the relief denudation and feed

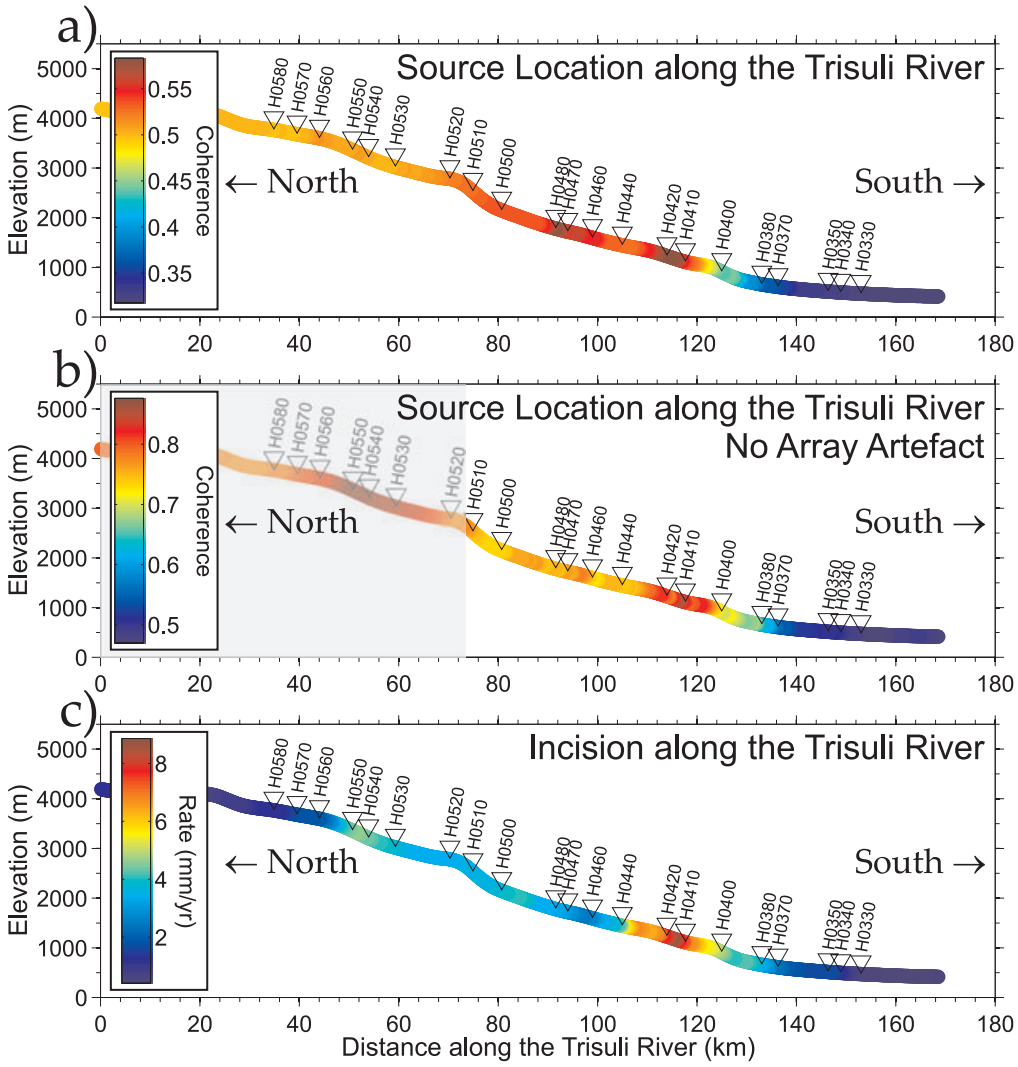

Figure 12. (a) Source location of the river seismic noise sources along the elevation profile of the stream. Red colours stand for areas of strong coherence (most active segment of river), whereas blue colours stand for areas of low coherence. (b) Source location of the river seismic noise sources corrected from the array artefact, using the results of the simulation S4 (Fig. 10). The grey-shaded area represents the section of the river where results are poorly constrained and discarded from the discussion. (c) River incision rate along the stream from Lavé & Avouac (2001). Red and blue colours stand for high and low incision rates, respectively.

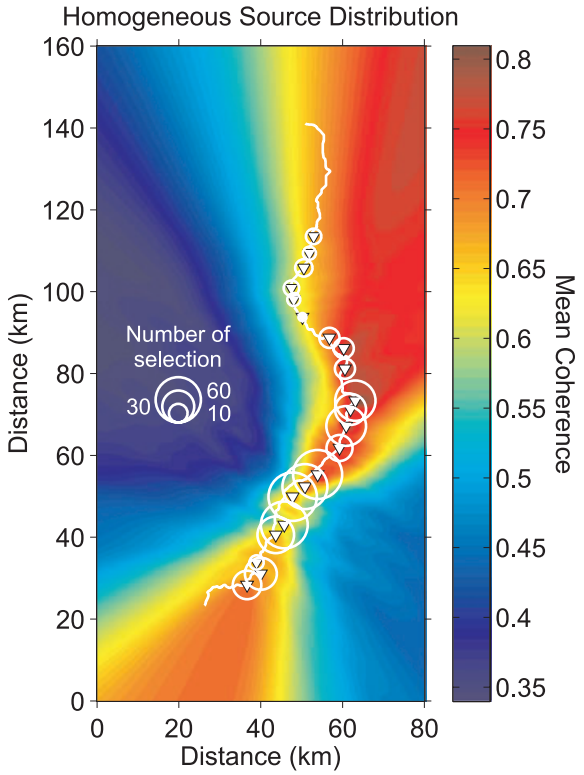

Figure 13. Coherence map for S4 migration (homogeneous source distribution). For the reliable set of frequencies (see text for precision), we average the best coherence maps. We then estimate from this mean solution the array artefact along the Trisuli River that we use for the correction in Fig. 12b. Red and blue colours stand for areas of strong and low coherence, respectively. Inverse white triangles represent the processed Hi-CLIMB stations. The size of white circles above each station location represents the number of times a station is selected. The selection of northernmost stations is reduced with regard to the portion H0400–H0480.

the Trisuli River with sediments that are tools to produce seismic noise.

In the High Range, north of station H0440, the decrease in the amount of seismic sources is lower than the decrease in erosion rate inferred by Lavé & Avouac (2001) (Fig. 12). The variation in the lithology of the bedrock along the river may explain this discrepancy. Station H0410 is located in the Lesser Himalaya (LH) that consists of metasediment whereas the stations north of H0440 are in the Higher Himalaya Crystalline (HHC) unit, consisting mainly of gneisses and quartzites (Upreti 1999). Erosion mechanisms are partly controlled by the lithology through the coefficient of erodability, denoting the rock strength (Sklar & Dietrich 2001). Lavé & Avouac (2001) assumed a lower coefficient for the HHC unit than the LH. Therefore, the bedrock in the HHC would tend to be more resistant to abrasion than in the LH. Because we monitor impacts of bed load instead of erosion rate, seismic noise may indicate that the sediment load north of H0440 is still large whereas the incision rate is low due to more resistant bedrock.

6 CONCLUSIONS

Our study confirms the potential of using NCFs to locate sources of high-frequency seismic noise when receivers are nearby. We use this technique to locate the segments of the Trisuli River responsible for the increase of seismic noise observed on a subset of Hi-CLIMB stations. Although the geometry of the network was not designed for this purpose, we show that most seismic sources are located along the steepest portion of the river, from H0410 and H0480. We even notice an overall good agreement between these sources from sediment transport and the incision rates deduced from other geomorphic approaches. According to the studied method, we may have an opportunity to monitor the bed load transport and potentially to deduce bedrock erosion.

The ability to recover the location of the seismic noise sources is due to the high density of Hi-CLIMB stations and the large variation of seismic noise in the monsoon season. Both factors help to isolate in space, time and frequency the seismic signal which effectively originates from the river. The performed synthetic tests confirm that the linear shape of the Hi-CLIMB network is not dedicated for the purpose of our study and produce some biases, especially on the northern part of the profile. However, these tests indicate that in the central part of the network, our results are reliable. We also inferred from the simulations that the correlation between seismic signals is not induced by *P* waves and rather reflects the coherence of *S* waves and/or Rayleigh surface waves.

A more suitable network would consist of stations deployed in 2-D, surrounding the targeted river. However, the lateral extent of such a dedicated network should remains small because the high-frequency signal generated by bed load transport in rivers rapidly attenuates. In addition, the background seismic noise, unrelated to the river as the anthropogenic perturbations, has to remain low. Otherwise, a selection stage has to be performed to isolate the river seismic noise. Despite such limitations, this approach seems to offer a great potential to continuously and spatially monitor the sediment transport in rivers without performing hazardous measurements in the stream. Finally, if a calibration is made between the lithology and the seismic signal produced by particles impacts, we may expect to seismically estimate river incision rate.

ACKNOWLEDGMENTS

The authors thank J. Nábělek and the entire Hi-CLIMB team for the fieldwork and data acquisition, in particular S. Sapkota and his colleagues from the Department of Mines and Geology of Nepal. The authors thank R. Hermann of Saint Louis University, St. Louis, Missouri, for providing the software package used to compute the synthetic seismograms. The authors thank the two anonymous reviewers for their valuable comments and suggestions on the paper. Project Hi-CLIMB is supported by the U.S. NSF Continental Dynamics Program, EAR 9909609.

REFERENCES

Bensen, G.D., Ritzwoller, M.H., Barmin, M.P., Levshin, A.L., Lin, F., Moschetti, M.P., Shapiro, N.M. & Yang, Y., 2007. Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements, *Geophys. J. Int.,* **169,** 1239–1260.

Bromirski, P.D. & Duennebier, F.K., 2002. The near-coastal microseism spectrum: spatial and temporal wave climate relationships, *J. geophys. Res.,* **107**(B8), 2166, doi:10.1029/2001JB000265.

Brzak, K., Gu, Y.J., Ökeler, A., Steckler, M. & Lerner-Lam, A., 2009. Migration imaging and forward modeling of microseismic noise sources near southern Italy, *Geochem. Geophys. Geosyst.,* **10**, doi:10.1029/2008GC002234.

Burtin, A., Bollinger, L., Vergne, J., Cattin, R. & Nábělek, 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**, doi:10.1029/2007JB005034.

Burtin, A., Bollinger, L., Cattin, R., Vergne, J. & Nábělek, J.L., 2009. Spatiotemporal sequence of Himalayan debris flow from analysis of high-frequency seismic noise, *J. geophys. Res.,* **114**.

Gabet, E.J., Burbank, D.W., Pratt-Sitaula, B., Putkonen, J. & Bookhagen, B., 2008. Modern erosion rates in the High Himalayas of Nepal, *Earth planet. Sci. Lett.,* **267**(3–4), 482–494.

Gutenberg, B., 1958. Microseisms, *Adv. Geophys.,* **5,** 53–92.

Herrmann, R.B., 2002. *An Overview of Synthetics Seismogram Computation: Computer Programs in Seismology, Version 3.30,* chap. Wavenumber Integration, pp. 3–1 3–13.

Hetényi, G., 2007. Evolution of deformation of the Himalayan prism: from imaging to modelling, *PhD thesis*, École Normale Supérieure - Université Paris-Sud XI, 400 pp.

Howard, A.D., Dietrich, W.E. & Seidl, M.A., 1994. Modeling fluvial erosion on regional to continental scales, *J. geophys. Res.,* **99**(B7), 13 971–13 986.

Kedar, S., Longuet-Higgins, M., Webb, F., Graham, N., Clayton, R. & Jones, C., 2008. The origin of deep ocean microseisms in the North Atlantic Ocean, *R. Soc. Lond. Proc. Ser. A,* **464,** 777–793.

Lavé, J. & Avouac, J.-P., 2001. Fluvial incision and tectonic uplift across the Himalayas of central Nepal, *J. geophys. Res.,* **106**(B11), 26 561–26 592.

Longuet-Higgins, M.S., 1950. A theory of the origin of microseisms, *Phil. Trans. R. Soc.,* **243,** 1–35.

Marston, R.A., Miller, M.M. & Devkota, L.P., 1998. Geoecology and mass movement in the Manaslu-Ganesh and Langtang-Jugal Himals, Nepal, *Geomorphology,* **26**(1–3), 139–150.

McNamara, D.E. & Buland, R.P., 2004. Ambient Noise Levels in the Continental United States, *Bull. seism. Soc. Am.,* **94**(4), 1517–1527.

Mohorovičić, A., 1916. Die Bestimmung des Epizentrums eines Nahbebens, *Gerlands Beitrage zur Geophysik,* **14,** 199–205.

Monsalve, G., Sheehan, A., Schulte-Pelkum, V., Rajaure, S., Pandey, M.R. & Wu, F., 2006. Seismicity and one-dimensional velocity structure of the Himalayan collision zone: earthquakes in the crust and upper mantle, *J. geophys. Res.,* **111**, doi:10.1029/2005JB004062.

Nábělek, J. *et al*., 2009. Underplating in the Himalaya-Tibet collision zone revealed by the Hi-CLIMB experiment, *Science,* **325**(5946), 1371–1374.

Pedersen, H.A., Krger, F. & the SVEKALAPKO Seismic Tomography Working Group, 2007. Influence of the seismic noise characteristics on noise correlations in the Baltic shield, *Geophys. J. Int.,* **168,** 197–210.

Putkonen, J., 2004. Continuous snow and rain data at 500 to 4400 m altitude near Annapurna, Nepal, 1999-2001, *Arct. Antarct. Alp. Res.,* **36,** 244–248.

Rhie, J. & Romanowicz, B., 2004. Excitation of Earth's continuous free oscillations by atmosphere-ocean-seafloor coupling, *Nature,* **431,** 552–556.

Rhie, J. & Romanowicz, B., 2006. A study of the relation between ocean storms and the Earth's hum, *Geochem. Geophys. Geosyst.,* **7**, doi:10.1029/2006GC001274.

Rost, S. & Thomas, C., 2002. Array seismology: methods and applications, *Rev. Geophys.,* **40**(3), 1008, doi:10.1029/2000RG000100.

Roux, P., Sabra, K.G., Gerstoft, P., Kuperman, W.A. & Fehler, M.C., 2005a. P-waves from cross-correlation of seismic noise, *Geophys. Res. Lett.,* **32**, doi:10.1029/2005GL023803.

Roux, P., Sabra, K.G., Kuperman, W.A. & Roux, A., 2005b. Ambient noise cross correlation in free space: theoretical approach, *J. acoust. Soc. Am.,* **117**(1), 79–84.

Sabra, K.G., Gerstoft, P., Roux, P., Kuperman, W.A. & Fehler, M.C., 2005. Surface wave tomography from microseisms in southern California, *Geophys. Res. Lett.,* **32**, doi:10.1029/2005GL023155.

Shapiro, N.M. & Campillo, M., 2004. Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise, *Geophys. Res. Lett.,* **31**, doi:10.1029/2004GL019491.

Shapiro, N.M., Ritzwoller, M.H. & Bensen, G.D., 2006. Source location of the 26 sec microseism from cross-correlations of ambient seismic noise, *Geophys. Res. Lett.,* **33**, doi:10.1029/2006GL027010.

Sklar, L.S. & Dietrich, W.E., 2001. Sediment and rock strength controls on river incision into bedrock, *Geology,* **29,** 1087–1090.

Snieder, R., 2004. Extracting the Green(tm)s function from the correlation of coda waves: a derivation based on stationary phase, *Phys. Rev. E,* **69,** 046610, doi:10.1103/PhysRevE.69.046610.

Stehly, L., Campillo, M. & Shapiro, N.M., 2006. A study of the seismic noise from its long-range correlation properties, *J. geophys. Res.,* **111**, doi:10.1029/2005JB004237.

Thouret, J.C., 1983. *Géodynamique des grands versants de l'Ankhu Khola, Népal central,* chap. Field research on slopes evolution in the Himalaya range: the example of the Ankhu Khola valley, Central Nepal, p. 231, CNRS edn.

Upreti, B.N., 1999. An overview of the stratigraphy and tectonics of the Nepal Himalaya, *J. Asian Earth Sci.,* **17**(5), 577–606.

Weaver, R.L. & Lobkis, O.I., 2001. Ultrasonics without a source: thermal fluctuation correlations at MHz frequencies, *Phys. Rev. Lett.,* **87**(13), 134301, doi:10.1103/PhysRevLett.87.134301.

Whipple, K.X. & Tucker, G.E., 1999. Dynamics of the stream-power river incision model: implications for height limits of mountain ranges, landscape response timescales, and research needs, *J. geophys. Res.,* **104**(B8), 17 661–17 674.

Yang, Y. & Ritzwoller, M.H., 2008. Characteristics of ambient seismic noise as a source for surface wave tomography, *Geochem. Geophys. Geosyst.,* **9**, doi:10.1029/2007GC001814.

Yao, H., van der Hilst, R.D. & de Hoop, M.V., 2006. Surface-wave array tomography in SE Tibet from ambient seismic noise and two-station analysis - I. Phase velocity maps, *Geophys. J. Int.,* **166,** 732–744.