

Modeling the foreshock sequence prior to the 2011, M_W 9.0 Tohoku, Japan, earthquake

D. Marsan¹, and B. Enescu²

Abstract. The 2011 M_W 9.0 Tohoku earthquake, Japan, was preceded by a 2 day-long foreshock sequence, initiated by a M_W 7.3 earthquake. We analyze this foreshock sequence, with the aim of detecting possible aseismic deformation transients that could have driven its evolution. Continuous broad-band recordings at F-net stations are processed to identify as exhaustive a set of $m_{JMA} > 1.2$ earthquakes as possible. We moreover directly quantify with these recordings the changes in detection level associated with changes in seismic or environmental noise. This earthquake dataset is then modeled, to show that the whole sequence can be readily explained without the need to invoke aseismic transients. The observation of a 3-hour long low-frequency noise increase, concurrent with an apparent migration of seismicity towards the epicenter of the impending M_W 9.0 megathrust earthquake, however suggests that some premonitory slip could have played a role in loading the asperity which failure initiated the M_W 9.0 shock. We thus propose that this aseismic slip, if it really existed, had only a minor role in triggering and southward displacing the foreshock sequence, as compared to earthquake interaction mechanisms that allow earthquakes to trigger one another.

1. Introduction

The observation of foreshocks prior to large earthquakes has raised for many years the hope that earthquake prediction could, at least to some extent, be approached by examining the time evolution of seismic activity (Jones and Molnar, 1979), or by searching for specific patterns of seismicity (Mogi, 1979). Stacking of pre-seismic activity over many mainshocks reveal an acceleration of the occurrence rate of foreshocks, in the vicinity of the impending mainshock (e.g., Maeda, 1999). This average pattern has moreover been found to be more frequent prior to thrust earthquakes in subduction zones as compared to, for example, strike-slip mainshocks (Reasenber, 1999), suggesting specific mechanisms at play during these sequences. Among those mechanisms, precursory slip on the fault at the mainshock hypocenter has been proposed, especially as large amounts of afterslip typically follow subduction zone earthquakes, indicating only partial coupling of the interface, hence the possibility that aseismic slip episodes could co-exist with 'normal' earthquakes in such zones. In this scenario, slow slip would then generate both the foreshocks and the mainshock by loading local asperities.

An alternative view (Helmstetter and Sornette, 2003; Felzer et al., 2004) is that foreshock sequences are caused by the same interaction processes that lead to aftershock triggering, and more generally that are involved in earthquake triggering by previous earthquakes, for example through stress transfer. Then, the acceleration of foreshock rates as one gets closer in time to the mainshock emerges only because the probability of occurrence of a second earthquake

increases immediately after the occurrence of a first. According to this model, no premonitory slip is required anymore to explain the accelerating pattern. Since the magnitude of the impending mainshock is independent of the pre-shock pattern (see also Abercrombie and Mori, 1996), the prediction prospects related to foreshock sequences are thus drastically reduced.

In order to test these two conflicting views, the analysis and modelling of individual foreshock sequences is required. We here present such a study, for the 3/11/2011, M_W 9.0 earthquake that occurred at the subducting interface offshore Tohoku in Japan. This earthquake was indeed preceded by a foreshock sequence, initiated about 50 km from the mainshock epicenter by a M_W 7.3 earthquake that occurred 2 days and 3 hours before the M_W 9.0 event. The geographical and temporal proximity of the two shocks is unlikely to be due to pure luck. However, it is unclear how the two events are related to each other. They could both result from a common mechanism, e.g., large-scale slip of the interface, possibly accelerating at the end of a preparatory process, and / or the first shock and its aftershocks could have further loaded the asperity that initiated the M_W 9.0 earthquake, hastening its occurrence.

Figure 1 displays the locations and times of occurrence of all $m_{JMA} \geq 3.5$ earthquakes listed in the Japanese Meteorological Agency (JMA) dataset, that occurred between the 3/9/2011 00:00 and the M_W 9.0 Tohoku earthquake, in an extended zone comprising the epicenters of the two shocks. A particularly interesting feature is the existence of four $m_{JMA} > 6$ earthquakes occurring in an interval of 3 hours, culminating in a $m_{JMA} = 6.8$ event that makes it the second biggest quake of this sequence. These earthquakes occur relatively late in the sequence (15 to 18 hours after the M_W 7.3 mainshock), and, perhaps more interestingly, they appear to be slightly off the rupture zone as highlighted by the previous aftershocks, going towards the epicenter of the upcoming M_W 9.0 event, see Figure 2. This migration of the seismic activity could indeed suggest a slow deformation migrating transient prior to the destructive shock (Ando and Imanishi, 2011). It has been suggested that the

¹ISTerre, Université de Savoie, CNRS, Le Bourget du Lac, France.

²Earthquake and Volcano Research Unit, National Research Institute for Earth Science and Disaster Prevention, Tsukuba, Japan.

migration pattern could have effectively started one month earlier, during an intermittent activity with relatively low earthquake rates, characterized by a maximum magnitude of $m_{JMA} = 5.5$ (Hirose et al., 2011).

The goal of this work is to investigate this foreshock sequence, in order to find arguments for or against the existence of such a slow transient. To do so, we will analyze the distribution, in space, time and magnitude, of the earthquakes belonging to this sequence. We make use of the JMA earthquake catalogue, but complement it with our own set of events as deduced directly from continuous broad-band recordings of the sequence by F-net stations (Okada et al., 2004). This allows us to track the changes in detection level, and more importantly to account for them, which is particularly important when analyzing such a short sequence.

2. Earthquake detection

2.1. Envelope

We use the vertical channel, sampled at 20 Hz, of the four F-net broad-band stations closest to the $M_W7.3$ earthquake, namely from north to south: TYS, KSN, KSK, and HRO, see Figure 1. We are interested in detecting as many aftershocks as possible, which can go undetected as they are masked in the coda wavetrain of larger earthquakes. To reduce this effect, we filter these signals with a 20-Hz high-pass filter. The envelopes are then computed, and shifted back in time by 28.3s, 25.4s, 34.8s and 35.6s for TYS, KSN, KSK and HRO, respectively. These time shifts correspond to the mean propagation time of P waves for earthquakes originating in the rupture zone of the $M_W7.3$ earthquake, as empirically derived by comparing the origin times (as given by the JMA catalog) and the observed arrival times for 5 large ($m_{JMA} > 6$ earthquakes) that occurred during this sequence. These 5 events include the $M_W7.3$ mainshock; they were selected because the P-arrival times can be determined with very good accuracy at the 4 stations. The envelopes are finally averaged together, and smoothed with a 100s low-pass filter, to give the smoothed, averaged envelope $S(t)$.

2.2. Detection

We detect earthquakes by searching when the envelope $S(t)$ increases by more than 10% over an interval of 30s. This duration is close to the $\simeq 40$ s P-S time for the earthquakes originating in this zone as seen by the 4 stations. Every time this increase is observed, we identify a new earthquake, with an occurrence time T corresponding to the start of the increase, and a maximum log-amplitude $M' = \log S$ computed as the first maximum of $S(t)$ after T . The choice of the parameters for the detection algorithm is done on visual inspections and a trial-and-error approach, by checking that all large JMA earthquakes are correctly detected and characterized, and that the S-wave or converted phases arrivals did not trigger spurious detections. Furthermore, we characterize the noise level prior to each detected event by $M'_0 = \log S(T)$. This selection finds 1504 earthquakes between 3/9/2011 00:00 and 3/11/2011 05:46, at the time of the $M_W9.0$ mainshock.

2.3. Magnitudes

In order to determine the equivalent JMA magnitudes of these earthquakes, we search in the JMA catalog the events that occurred in the $[38^\circ, 39^\circ]$ latitude and $[142.2^\circ, 143.7^\circ]$ longitude zone centered on the $M_W7.3$ mainshock, during these 2.24 days. We found 245 such events. Each one of these is then compared to our detected earthquakes; namely, we check whether there exists one of our 1504 events with

an occurrence time T within 10s of the JMA earthquake. We find 200 such pairs. A plot of m_{JMA} vs. M' for these 200 events shows a clear linear relation, see Figure 3: $m_{JMA} = 0.6212M' + 2.0204$. Only 9 out of these 200 earthquakes are more than 0.5 units of m_{JMA} away from this trend. Out of the 245 JMA earthquakes, 45 are not associated to one of our 1504 events. Visual inspection shows that 36 out of 45 have time differences with the closest new detected earthquake greater than 10s (up to 48s), see Figure 4. We deduce from this comparison that the uncertainty on T is equal to 11.5s. The 9 remaining JMA earthquakes not paired with one of our 1504 events occur within less than 60s from another one of the 245 JMA earthquakes, and for these 9 doublets our routine identifies only one event, but with a maximum log-amplitude equal to that of the greatest of the two events. We therefore compute M , an equivalent JMA magnitude for all 1504 earthquakes, as $M = 0.6212M' + 2.0204$. The minimum magnitude is $M = 0.55$, while the maximum is 7.43 and corresponds to the $M_W7.3$ mainshock. Equivalently, the noise magnitude M_0 is computed as $M_0 = 0.6212M'_0 + 2.0204$, and the envelope $S(t)$ is also translated into a JMA-magnitude equivalent envelope with $\mu(t) = 0.6212 \log S(t) + 2.0204$.

2.4. Geographical origin

Of the 1504 identified events, some could correspond to earthquakes not located within the zone of interest. Since we do not locate the earthquakes, we here only estimate how significant this problem is. To do so, we select all the earthquakes listed in the JMA catalog for the 2.24 day-long period of interest (starting at 00:00 on the 3/9/2011), and separate them into two populations: (1) those within 100 km of the $M_W7.3$ mainshock epicenter, and (2) those located more than 100 km away. Plotting in Figure 5 the number of earthquakes vs. their JMA magnitude for the two populations, we see that the first population has a significantly higher detection threshold, the magnitude of completeness being estimated to 4.2, while it is equal to 2.1 for the second population. Extrapolating the Gutenberg-Richter law below these magnitudes of completeness, we find that, for all magnitude bands, there are 17 times more earthquakes in the first population than in the second. We thus conclude that no more than $1/18 = 5.5\%$ of the 1504 detected earthquakes could actually be located outside the zone of interest. Given that the second population includes earthquakes that can be very far from the four stations, and that thus can go undetected by them (especially as we only look at frequencies above 20 Hz), we further argue that this proportion is likely over-estimated.

2.5. Changes in minimum noise level

Figure 6 shows the JMA magnitude-equivalent envelope $\mu(t)$, along with the 1504 detected events at times T and with magnitudes M . Fluctuations in the minimum noise level are observed, that are likely related to anthropogenic noise at the high frequencies probed by our processing ($f > 20$ Hz). To avoid contamination of our analysis by such a spurious effect, we impose a cut-off magnitude equal to 1.2, see Figure 6. In the following, we will thus only consider the 979 detected earthquakes with $M > 1.2$.

As an illustration, we show in Figure 7 a 2 hour-long window containing the $m_{JMA} = 6.8$ shock. All JMA earthquakes in the zone of interest are well picked and characterized by our method; the error on the magnitude estimate is small, of the order ± 0.1 , as expected from Figure 3. There are several other JMA earthquakes, that are located outside the zone, but that do not correspond to a peak in the envelope, and are thus correctly skipped by our selection method.

3. Accounting for undetected earthquakes

Small earthquakes can go unlisted in earthquake catalogues because they occur too shortly after a larger shock, so that the waves they radiate are hidden, especially in the coda, or interfere with independent wavetrains originating at nearly coincident sources, so that the arrival times of the P and / or S phases cannot be accurately picked. As a consequence, the detectability of small earthquake changes over time. This has to be accounted for when monitoring the evolution of the seismic activity. This is a two-step process: (1) the detectability must be quantified through time, and then (2) the seismicity model to be fitted to the data must account for this detectability. The first step can be particularly problematic when analyzing earthquake catalogues for which the processing chain is not fully known and / or the raw seismic waveforms are not available. Our method detailed above for detecting earthquakes in the proximity of the $M_W 7.3$ shock has the advantage that we can thoroughly characterize the time evolution of the detectability, in a way initially inspired by the treatment of Peng et al. (2007).

3.1. Detectability

Knowing the magnitude M of the earthquakes and the magnitude M_0 of the noise immediately prior to them, we analyze the distribution of the difference $\delta M = M - M_0$ in Figure 8. It follows a Gutenberg-Richter law with a sharp cut-off at $\delta M = 0$, so that no earthquake is detected with a magnitude smaller than the current level of noise. The b -value equals 0.47, which is within the uncertainty of the b -value of the JMA dataset (0.53 ± 0.11). We therefore model the detectability with a simple cut-off in magnitude at $\mu(t)$: if an earthquake of magnitude m occur at time t , then it is detected with certainty if $m > \mu(t)$, while it is missed with certainty if $m < \mu(t)$.

3.2. Modelling seismicity with time-varying detectability

We model earthquake occurrence as a non-homogeneous Poisson model with rate $\lambda(t)$. For any time interval $[t_a, t_b]$, the number of earthquakes - both detected and undetected - with magnitude greater than the cut-off magnitude 1.2 is a realization of a Poisson law with mean $\Lambda = \int_{t_a}^{t_b} dt \lambda(t)$. Denoting by $\pi(t)$ the probability that an earthquake of magnitude greater than 1.2 occurring at time t is effectively detected, it can be shown that the cost function to be minimized is

$$J = \int_0^T dt \pi(t) \lambda(t) - \sum_i \log \lambda(T_i) \quad (1)$$

where T_i are the occurrence times of the detected earthquakes, and $[0, T]$ is the time span of the dataset. The Poisson mean Λ_d of detected earthquakes in any time interval $[t_a, t_b]$ is $\Lambda_d = \bar{\pi} \Lambda$, where $\bar{\pi} = \frac{1}{t_b - t_a} \int_{t_a}^{t_b} dt \pi(t)$ is the average of $\pi(t)$ for this time interval. For intervals $[T_n, T_{n+1}]$ between two consecutive detected earthquakes with $M > 1.2$, and assuming that the rate $\lambda(t)$ is piecewise constant over these intervals, i.e., $\lambda(T_1 < t < T_2) = \lambda_1$, $\lambda(T_2 < t < T_3) = \lambda_2$, and so on, the maximum likelihood estimate of Λ_d for any interval is simply 1. The corresponding Poisson mean for all (detected and undetected) earthquakes is then $\frac{1}{\bar{\pi}}$, and the MLE rate is therefore $\lambda(T_n < t < T_{n+1}) = \frac{1}{\bar{\pi}(T_{n+1} - T_n)}$.

As explained above, the detectability is a simple cut-off at magnitude $\mu(t)$. The probability $\pi(t)$ of detecting an earthquake occurring at t with magnitude greater than 1.2

is thus equal to 1 if $\mu(t) < 1.2$, or equal to $10^{-b(\mu(t)-1.2)}$ if $\mu(t) > 1.2$, with $b = 0.47$ (Figure 8). We compute $\pi(t)$ from $\mu(t)$, and deduces the rate $\lambda(t)$. Figure 9 displays $\lambda(t)$ and its integral over time, giving the Poisson mean of the cumulative number of (detected and undetected) earthquakes with $M > 1.2$. We obtain a low p -value of 0.51. As a comparison, estimating p with no correction for changes in detection but excluding the first hour after the $M_W 7.3$ event, and for a varying magnitude cut-off ranging from 3 to 5, yield $0.6 < p < 0.8$.

3.3. Testing the correction

We test this method by simulating earthquake sequences with characteristics similar to those of the studied sequence. By construction, we know the real number of total earthquakes for these simulations, and we can therefore analyze whether the proposed correction is able to accurately estimate this number. The synthetic sequences are generated with an ETAS model (Kagan and Knopoff, 1981; Ogata, 1988) with no background rate, so that the modeled rate of earthquake at time t is $\lambda(t) = \sum_{n/T_n < t} A e^{\alpha M_n} (t + c - T_n)^{-p}$

where T and M are the time of occurrence and magnitudes of the simulated earthquakes, and A , α , p and c are model parameters. We here take $p = 1.1$, $c = 10$ s, $\alpha = 1.037$, and $A = 0.0209$. The magnitudes are distributed according to a Gutenberg-Richter law with b -value equal to 0.5, with a lower bound at 0 and an upper bound at 7.3. With this choice of parameters, we have that $\alpha = 0.9 \times b \times \log 10$, hence $\alpha < b \times \log 10$, i.e., small earthquakes slightly dominate big earthquakes in terms of collectively triggering new earthquakes. Moreover, the branching ratio is 2.5. This ratio measures the mean number of triggered aftershocks per mainshock, unconditioned on the magnitude of the latter. The fact that it is greater than 1 is not a problem in this application, since we limit the duration of the sequence to just 2 days.

The sequence is initiated with a mainshock of magnitude 7.3 at $t = 0$, and the model is run to generate earthquakes up to 2 days after this mainshock. From this time series, we simulate the envelope $\mu(t)$ with the following model:

- an earthquake at time T and with magnitude M is associated to an envelope (in linear, not logarithmic, scale) $s(t < T) = 0$ and $s(t > T) = e^{M \times f(t-T)}$ with $f(t) = \frac{t}{\tau} e^{\frac{1}{a}[1-(t/\tau)^a]}$. This choice of $f(t)$ is done on empirical grounds, as it gives a good fit to the envelope $S(t)$ found for the real sequence, providing that the time τ giving the maximum of f is set to 40 s, and $a = 1/2$.

- The total envelope is then computed from the individual envelopes $s_n(t)$, where n is the index of the considered earthquake, as $S^q(t) = \sum s_n^q(t)$. Parameter q characterizes the non-linearity of this 'stacking'. For $n \gg 1$, $S(t)$ tends to $\max_n s_n(t)$. We empirically find that $q = 6$ provides a good fit to the envelope of the real sequence.

- The (logarithmic) envelope is finally computed as $\mu(t) = \log S(t)$.

Applying this to the detected earthquakes of the real sequence, we find that the modeled envelope mimics the real envelope with very good accuracy, with the notable exception of the coda of the largest ($M > 6$) shocks, see Figure 10.

For any given synthetic catalogue, we thus generate the envelope $S(t)$ and the associated 'magnitude' $\mu(t)$. Then,

the detection procedure described in Section 2 is applied. This yields a set of detected earthquakes with occurrence times T and estimated magnitudes M that can differ from the true occurrence times and magnitudes of the simulated earthquakes. In particular, we recall that the standard deviation of the error in occurrence time is 11.5s for the real data; for the synthetics, this standard deviation is 18s. As with the real sequence, we select all the detected earthquakes with $M > 1.2$, and compute the corrected Poisson average $\Lambda(T_n, T_{n+1})$ between two consecutive detected earthquakes.

Two hundreds independent synthetic datasets were thus generated, allowing us to compare an ensemble averaged corrected rate $\lambda(t)$ to the true rate of $M > 1.2$ earthquakes. Figure 11 shows this comparison. The corrected rate is very close to the true rate, on ensemble average. Only for the first two minutes after the $M = 7.3$ mainshock does the corrected rate under-estimate the true rate, possibly owing to a lack of detected earthquakes in this time interval, which forces the estimated rate to be constant (as already observed and explained in Figure 9). We thus conclude that the method is well able to estimate the rate of occurrence of earthquakes, even for magnitudes that can be temporary affected by incomplete detection.

4. Searching for aseismic deformation using seismicity data

We now investigate whether aseismic slip could have played a role in triggering this two day-long sequence and the subsequent $M_W 9.0$ mainshock, especially given the observed southward migration evidenced by the locations of three of the four $m_{JMA} > 6.0$ shocks after 15 to 18 hours following the $M_W 7.3$ earthquake. Previous works on similar issues, but for other mainshocks, were largely based on stress calculations: shear stress imparted by the foreshocks on the hypocenter of the impending $M_W 7.3$, 1975 Haicheng earthquake (Jones et al., 1982), Coulomb stress generated by earlier foreshocks on subsequent foreshocks of the $M_W 7.3$, 1992 Landers earthquake (Dodge et al., 1995), and Coulomb stress caused by foreshocks on a set of California mainshocks (Dodge et al., 1996) or at the hypocenter of a moderate-size thrust earthquake in Japan (Umino et al., 2002). In all cases but one (the Mount Lewis earthquake, discussed in Dodge et al., 1996), the small amount or negative effect of the stress transferred argued in favor of premonitory aseismic slip. Our approach is complementary from these: we ask whether the time series of the foreshocks, rather than their relative locations, is coherent or not with a stochastic triggering model that includes aseismic forcing. More specifically, we test whether a foreshock at time t is either caused by the $M_W 7.3$ earthquake and by previous foreshocks at earlier times, or by a time varying aseismic forcing (e.g., pore fluid changes, aseismic slip, etc). Here, we only use the magnitude and occurrence times of the earthquakes.

We base our analysis on the seismicity dataset described in Section 2, and we account for the changes in detection probability as detailed in Section 3. Our treatment uses the fact that an episode of aseismic transient deformation can be detected as it generates an increase in earthquake activity, as has been evidenced in several instances (Sacks et al., 1981; Vidale and Shearer, 2006; Lohman and McGuire, 2007; Bourouis and Bernard, 2007; Llenos et al., 2009). Searching for such periods of increased activity is however made difficult by the natural fluctuations in earthquake rate observed during any aftershock sequence, caused by the triggering of aftershocks by previous aftershocks. In order to

discriminate between the two effects (loading by aseismic deformation, or loading by stress transfer from previous earthquakes), we describe the earthquake time series as resulting from the superposition of these two effects, and find the best parameter set that reproduce the observations.

We therefore model the detected earthquake time series using an ETAS model, accounting for the time-varying detectability probability $\pi(t)$ of $M > 1.2$ earthquakes. The model includes an aseismic forcing rate $\lambda_0(t)$, also called background rate, that can vary in time. The rate of occurrence of all (detected and undetected) earthquakes with magnitude $M > 1.2$ at time t is then modeled as

$$\lambda(t) = \lambda_0(t) + \sum_{n/T_n < t} A e^{\alpha(M_n - 1.2)} (t + c - T_n)^{-p} \quad (2)$$

The two terms on the right-hand side of Equation 2 are the two afore mentioned effects: $\lambda_0(t)$ represents the aseismic forcing process, while the sum represents the triggering of earthquakes by previous earthquakes. In the absence of such fault interactions, i.e., if earthquakes were unable to trigger aftershocks, then the observed rate of earthquakes would simply equal the forcing rate.

Methods have been developed for determining a time-varying forcing rate $\lambda_0(t)$ from earthquake time series (Daniel et al., 2011; Llenos and McGuire, 2011; Marsan et al., submitted). We here adapt the method described in Marsan et al. (submitted) to the present case. We first search for the best ETAS parameters A, α, p, c for a constant background rate $\lambda_0(t) = \lambda_0$. To do so, we minimize the cost function of Equation (1). We find that $A = 1.78 \times 10^{-6}$, $\alpha = 2.36$, $p = 0.49$, and $c = 0.1$ minute. The p -value is coherent with the one found for overall activity decay, see Figure 9. Parameter c is badly constrained; it serves as an artificial time cut-off to avoid divergence of the Omori-Utsu law. Given that our occurrence times have a $\simeq 10$ s resolution, we choose c to be of the order but less than this uncertainty, and therefore imposed $c = 0.1$ minute. We also performed similar analysis with $c = 1$ minute, and came to the same conclusions as with the $c = 0.1$ value. The best background rate is $\lambda_0 = 0.05$ per minute.

Given these parameters, we compute the probabilities ω_{ij} that earthquake i triggered earthquake j according to the optimized model, as well as the probabilities ω_{0j} that earthquake j is independent from all previous earthquakes in the dataset, i.e., is a background earthquake. More precisely:

$$\omega_{ij} = \frac{\lambda_{ij}}{\lambda(T_j)} \quad \text{and} \quad \omega_{0j} = \frac{\lambda_0(T_j)}{\lambda(T_j)} \quad (3)$$

where $\lambda_{ij} = A e^{\alpha(M_i - 1.2)} (T_j + c - T_i)^{-p}$ is the modeled rate of aftershocks of earthquake i at the time of occurrence of earthquake j . By construction, $\sum_{i < j} \omega_{ij} + \omega_{0j} = 1$. The time series made of the probabilities ω_{0j} is then smoothed in time, to yield the time-varying background rate $\lambda_0(t)$ given those probabilities. We here use a Gaussian filter with smoothing scale δt :

$$\lambda_0(t) = \sum_j \frac{\omega_{0j}}{\sqrt{2\pi} \delta t} e^{-(t - T_j)^2 / 2\delta t^2} \quad (4)$$

The algorithm then consists in iterating the two steps described above:

- given the parameters $\alpha = 2.36$, $p = 0.49$, $c = 0.1$ minute, an initial $A = 1.78 \times 10^{-6}$, and the background rate $\lambda_0(t)$ (initially constant, equal to 0.05 per minute), deduce the probabilities ω_{ij} , and the smoothed $\lambda_0(t)$ resulting from ω_{0j} ;

- optimize A given the other parameters and $\lambda_0(t)$; until convergence of both A and $\lambda_0(t)$ is obtained. While the initial choice of A and λ_0 does not influence the final solution, the choice of the smoothing parameter δt is critical: for a very short δt , typically much shorter than the mean of the waiting time between two successive detected earthquakes $T_{i+1} - T_i$, the background rate is equivalent to a Dirac comb, and the algorithm converges towards $A = 0$, hence all earthquakes are considered as background. In order to select δt , we run the algorithm for several values of this parameter, and compare the values of the cost function J of Equation (1) after convergence. Since a short δt allows for a background rate with rapid fluctuations, it must be penalized. We showed in Marsan et al. (submitted) that the Akaike Information Criterion $AIC = J + \frac{T}{\delta t}$ is a pertinent criterion to optimize δt .

Figure 12 displays the final background rate $\lambda_0(t)$ for three values of δt . The lowest AIC is found for $\delta t = 1000$ minutes, with $A = 2.09 \times 10^{-6}$. This implies an almost constant background rate $\lambda_0(t)$, with a mean rate of 0.0056 per minute, corresponding to only 26 of the 979 detected earthquakes seen as background earthquakes. For shorter δt , a transient increase in forcing is found for $1500 < t < 2000$ minutes, that relates to the occurrence of a mildly intense swarm of earthquakes with relatively low magnitudes, see Figure 6. However, since the overall sequence is most efficiently modeled by a nearly constant forcing, it suggests that this swarm is not anomalous and can be well explained by triggering through earthquake interactions. Moreover, it occurs after the four $m_{JMA} > 6$ earthquakes, and thus follows rather than precedes the observed southward earthquake migration.

Sensitivity studies show that this method can damp actual fluctuations in forcing rate, if the model parameters are badly estimated, more exactly if α or p are under-estimated and / or c is over-estimated. Since our $p = 0.49$ value is effectively low compared to typical p values, generally found in the 0.8 to 1.2 range (e.g., Hainzl and Marsan, 2008), we tested how a higher p -value would change the results. We thus imposed $p = 1$, keeping α and c the same as before, and performed the same analysis again. An optimal $\delta t = 100$ minutes is then obtained. The forcing $\mu(t)$ follows the same trend as the $\delta t = 100$ curve of Figure 12, although with much higher values, the mean being equal to 0.28 per minute, or equivalently 779 of the 979 detected earthquakes are found as background earthquakes. This appears to be an unrealistically high proportion of background events. Given that the optimized AIC is then larger by 29.8 units than the minimal AIC obtained with $p = 0.49$, we argue that assuming such a large p -value is inappropriate for this sequence.

We therefore conclude that no anomalous transient in forcing characterize the aftershock sequence of the $M_W 7.3$ shock, which is well modeled by an ETAS model with nearly constant background rate, see Figure 13, with only a small proportion ($< 3\%$) of earthquakes being due to non-seismic loading. While this does not preclude the existence of an aseismic transient, it shows that such a transient, if any, did not significantly affect the earthquake time series. More precisely, the estimated number of background earthquakes for the duration of the sequence corresponds, for our parameterized model, to the average number of aftershocks that a $m_{JMA} = 6.2$ shock would have triggered (if it had occurred at the beginning of the sequence). Hence aseismic loading is estimated to be less than a $m_{JMA} = 6.2$ event, and is found to be uniformly distributed in time.

5. Low-frequency noise anomaly

Inspecting the evolution in time of the recordings at frequencies lower than the 20 Hz cut-off imposed previously, we note an increase in noise for the 0.1 Hz - 1 Hz band, at F-net stations located the closest to the foreshock sequence. This increase is observed for at least the 3 hours immediately prior to the four $m_{JMA} > 6$ shocks, at 15 to 18 hours after the $M_W 7.3$ earthquake (purple curve in Figure 14, from $t = 880'$ to $t = 1070'$). The period characterized by this increase could potentially extend to $t > 1070'$, but the envelope of the coda wavetrains associated with the $m_{JMA} > 6$ shocks prevents us to determine its full extent.

We quantify this increase at each F-net station by computing $s(t)$, the 1000s-smoothed logarithm of the envelope of the 0.1 - 1 Hz band-filtered vertical recording. We then compute the ratio $R = \frac{\bar{s}_2 - \bar{s}_1}{\sigma}$, where $\bar{s}_{1,2}$ is the mean of $s(t)$ over the period 1 or 2 as shown in Figure 14, and σ is the standard deviation of $s(t)$ for the 24 hours on the 3/8/2011, hence prior to the foreshock sequence. This ratio R reaches values higher than 3 at KSK and HRO, that demonstrate a significant increase in noise, and positive values for F-net stations located close to the foreshock sequence, see Figure 15.

While the origin of this low-frequency noise anomaly is difficult to determine based on the sole F-net data, it is potentially related to the occurrence of the foreshock sequence. An increase in noise at frequencies lower than about 1.5 Hz has been observed in the 40 minutes prior to the $M_W 7.6$, 1999 Izmit earthquake (Bouchon et al., 2011), and could have been caused by pre-rupture slow slip at the mainshock asperity. Seismicity migrated during the $M_W 9.0$ Tohoku foreshock sequence, more particularly in relation to the occurrence of three of the four $m_{JMA} > 6$ shocks at 15 to 18 hours after the $M_W 7.3$ earthquake (see Figure 2). The low-frequency noise increase is observed immediately before these $m_{JMA} > 6$ shocks, and could therefore be related to slow slip at the southernmost end of the $M_W 7.3$ rupture. Such a slow slip could then have further loaded the $M_W 9.0$ asperity, leading to its catastrophic failure 1.5 days later. The existence of premonitory deformation transients lasting hours, and stopping with the occurrence of a large shock - here the burst of $m_{JMA} > 6$ earthquakes - has for example been documented by Melbourne and Webb (2002) in the case of a $M_W 7.6$ aftershock of the 2001, $M_W 8.4$ Peru earthquake.

GPS data reveal that two days of afterslip following the $M_W 7.3$ earthquake amounted to a $M_W 7.0$ event, with slip located mostly north and downdip from the hypocenter (Miyazaki et al., 2011). The noise in the GPS data does not allow to resolve a possible transient increase in slip rate concomitant with the seismic noise anomaly. Tiltmeter signals exhibit transient excursions during the first half of the foreshock sequence, but they appear incoherent from one station to the next (Hirose, 2011), so that local anelastic relaxation after the co-seismic shaking of the $M_W 7.3$ earthquake could be the cause.

6. Discussion and Conclusions

The two-day long foreshock sequence of the 2011 $M_W 9.0$ Tohoku earthquake provides a new opportunity to investigate how foreshock patterns develop, and whether any of

their characteristics could be related to the occurrence of the impending mainshock. It has become traditional to oppose two end member models. The first explains foreshocks as a cascade of ruptures that trigger one another. The second sees foreshocks as being caused by an aseismic process, in which case their occurrences are independent of each other.

Several tests of these models have been performed by checking whether static stress changes, either shear or Coulomb, are consistent with the development of the sequence and the eventual occurrence of the mainshock (Jones et al., 1982; Dodge et al., 1995, 1996; Umino et al., 2002). In a large majority of cases, these studies have refuted the model of cascading ruptures. Such calculations however suffer from many uncertainties, in hypocenter locations, rupture sizes, fault geometries, and friction coefficient, that require a probabilistic treatment (Dodge et al., 1996). More importantly, unloading of the main asperity should imply that the mainshock occurrence is delayed. As an example, the Landers mainshock was found to be unloaded by a few tenths of MPa (Dodge et al., 1996). According to the rate-and-state friction model (Dieterich, 1994), and assuming the foreshock sequence develops at the very end of the mainshock nucleation phase, this unloading $\Delta\tau < 0$ amounts to a time delay $\Delta t = T (1 - e^{-\Delta\tau/A\sigma})$ where T would have been the time to failure in the absence of a stress perturbation, and $A\sigma$ is a model parameter, here set to $A\sigma = 0.05$ MPa as proposed by Toda et al. (2005) in their modelling of the Landers aftershock sequence. Taking $\Delta\tau = -0.4$ MPa (cf. Figure 15 of Dodge et al., 1996) as a mean, representative value of the stress change, the time delay should be a few thousands times the unperturbed T . Since the foreshock sequence started 7 hours before the mainshock, this implies that the Landers mainshock would have occurred within less than 1 minute after the time of the first foreshock, if no foreshocks had perturbed its nucleation phase. According to this model, significant stress unloading by foreshocks is thus difficult to reconcile with the observation that they typically precede the mainshock by only a few hours to days.

Alternative tests of foreshock models can be performed, that do not require computing stress changes. Using a stochastic model of earthquake interactions, Felzer et al. (2002) argued that the 1999, $M_W 7.1$ Hector Mine earthquake, California, was triggered by foreshocks occurring within 1 day of the mainshock, which were themselves triggered by a chain of previous shocks going back to the 1992 Landers earthquake about 20 km away. Analyzing synthetic catalogs generated by a similar model, Helmstetter and Sornette (2003) demonstrated that commonly observed features of foreshock sequences are well reproduced without requiring the need for a preparatory phase to the mainshock that would generate foreshocks as a by-pass product.

We here adopt an approach that does not a priori exclude any of these two conflicting models. The foreshock sequence of the Tohoku earthquake is described as potentially resulting from both rupture cascading and aseismic loading, which rate is allowed to change over time so to simulate episodes of aseismic deformation. The latter ingredient is found to be negligible compared to the first, which by itself reproduces the observed time series in an efficient way.

The southward migration of seismicity observed at about 15 to 18 hours after the $M_W 7.3$ shock delineates a gap of about 15 to 20 km width, devoid of any earthquake, see Figure 2. This gap is likely to correspond to a creeping zone,

which underwent afterslip following the $M_W 7.3$ shock (Ando and Imanishi, 2011). After 15 hours, and possibly culminating with a 3 hour-long increase in slip rate, as suggested by the low frequency noise anomaly, loading by afterslip became eventually strong enough to trigger seismicity, and earthquake activity thus started on the south-west side of this gap. A burst of three $m_{JMA} > 6.0$ shocks then occurred, along with their own local aftershocks, resuming the cascade of ruptures at close distance to the $M_W 9.0$ asperity. Further loading of this asperity, possibly by both continuing afterslip and elastic stress transfer due to the $m_{JMA} > 6.0$ earthquakes and aftershocks, then led to the main rupture. In this scenario, both components (earthquake interactions and aseismic slip) participate to the development of the sequence. Since slow slip is only required to initiate the burst on the south-west side of the gap, it is not well resolved by our seismicity model.

It is worth noticing that the stress calculations and the stochastic modelling approach differ by their use, or not, of the spatial information. Negative loading of the main asperity by foreshocks is generally due to the co-location of one or several of the latter with the asperity. This co-location is also proposed in the case of the Izmit earthquake (Bouchon et al., 2001), although no precise location of the hypocenters was estimated. Negligible loading, as in the case of the Haicheng (Jones et al., 1982), the 1998 $M 5.0$ Sendai, Japan (Umino et al., 2002) or the 1990, $M_L 5.2$ Upland, California (Dodge et al., 1996) earthquakes is on the contrary due to the large distance separating the foreshocks and the mainshock, relative to the size of the foreshock ruptures. Unlike these calculations, which essentially depend on the relative locations of the events, the stochastic modelling of Felzer et al. (2002, 2004) and Helmstetter and Sornette (2003), or as described in this manuscript, is solely based on time and magnitude information. This clearly highlights the urgent need to further develop this type of modeling, to account for the locations of the earthquakes relative to one another. A trade-off between well resolved locations and an exhaustive detection of 'all' foreshocks above a low magnitude cut-off would then need to be found.

Acknowledgments

This work was supported by the French ANR ASEISMIC project. We would like to thank Olivier Lengliné, Michel Bouchon, Agnès Helmstetter, Pascal Bernard, François Jouanne, Philippe Lesage, Jean Vandemeulebrouck, Hitoshi Hirose and Takanori Matsuzawa for stimulating discussions on this topic.

References

- Abercrombie R., and J. Mori (1996), Occurrence patterns of foreshocks to large earthquakes in the western United States, *Nature*, **381**, 303–307.
- Ando R., and K. Imanishi (2011), Possibility of $M_W 9.0$ mainshock triggered by diffusional propagation of after-slip from $M_W 7.3$ foreshock, in press for *Earth Planets Space*.
- Bouchon M., H. Karabulut, M. Aktar, S. Özalaybey, J. Schmittbuhl, and M.-P. Bouin (2011), Extended Nucleation of the 1999 $M_W 7.6$ Izmit Earthquake, *Science*, **331**, 10.1126/science.1197341.
- Bourouis S., and P. Bernard (2007), Evidence for coupled seismic and aseismic fault slip during water injection in the

- geothermal site of Soultz (France), and implications for seismicogenic transients, *Geophys. J. Int.*, **169**, 723 – 732.
- Daniel G., D. Marsan, and M. Bouchon (2008), Earthquake triggering in southern Iceland following the June 2000 $M_s 6.6$ doublet, *J. Geophys. Res.*, **113**, 10.1029/2007JB005107.
- Daniel G., E. Prono, F. Renard, F. Thouvenot, S. Hainzl, D. Marsan, A. Helmstetter, P. Traversa, J. L. Got, L. Jenatton, and R. Guiguet (2011), Changes in effective stress during the 2003–2004 Ubaye seismic swarm, France, *J. Geophys. Res.*, **116**, 10.1029/2010JB007551.
- Dieterich, J. H. (1994), A constitutive law for rate of earthquake production and its application to earthquake clustering, *J. Geophys. Res.*, **99**, 2601–2618.
- Dodge D. A., G. C. Beroza, and W. L. Ellsworth (1995), Foreshock sequence of the 1992 Landers, California, earthquake and its implications for earthquake nucleation, *J. Geophys. Res.*, **100**, 9865–9880.
- Dodge D. A., G. C. Beroza, and W. L. Ellsworth (1996), Detailed observations of California foreshock sequences: Implications for the earthquake initiation process, *J. Geophys. Res.*, **101**, 22371–22392.
- Felzer K. R., T. W. Becker, R. E. Abercrombie, G. Ekström, and J. R. Rice (2002), Triggering of the 1999 $M_W 7.1$ Hector Mine earthquake by aftershocks of the 1992 $M_W 7.3$ Landers earthquake, *J. Geophys. Res.*, **107**, 10.1029/2001JB000911.
- Felzer K. R., R. E. Abercrombie, and G. Ekstrom (2004), A common origin for aftershocks, foreshocks, and multiplets, *Bull. Seismol. Soc. Am.*, **94**, 88–98.
- Hainzl S., and D. Marsan (2008), Dependence of the Omori p-value on mainshock magnitude in rate-and-state friction models, *J. Geophys. Res.*, **113**, 10.1029/2007JB005492.
- Helmstetter A., and D. Sornette (2003), Foreshocks explained by cascades of triggered seismicity, *J. Geophys. Res.*, **108**, 2457.
- Hirose H. (2011), Tilt records prior to the 2011 off the Pacific coast of Tohoku earthquake, *Earth Planets Space*, in press.
- Hirose F., K. Miyaoka, N. Hayashimoto, T. Yamazaki, and M. Nakamura (2001), Outline of the 2011 off the Pacific coast of Tohoku earthquake ($M_W 9.0$) - seismicity: foreshocks, mainshock, aftershocks and induced activity, *Earth Planets Space*, in press.
- Jones L. M., B. Wang, S. Xu, and T. J. Fitch (1982), The foreshock sequence of the February 4, 1975, Haicheng earthquake ($M=7.3$), *J. Geophys. Res.*, **87**, 4575–4584.
- Jones L. M., and P. Molnar (1979), Some characteristics of foreshocks and their possible relationship to earthquake prediction and premonitory slip on faults, *J. Geophys. Res.*, **84**, 3596–3608.
- Kagan Y. Y., and L. Knopoff (1981), Stochastic synthesis of earthquake catalogs, *J. Geophys. Res.*, **86**, 2853–2862.
- Llenos A. L., J. J. McGuire, and Y. Ogata (2009), Modeling seismic swarms triggered by aseismic transients, *Earth Plan. Sc. Lett.*, **281**, 59 – 69.
- Llenos A. L., and J. J. McGuire (2011), Detecting aseismic strain transients from seismicity data, *J. Geophys. Res.*, **116**, 10.1029/2010JB007537.
- Lohman R. B., and J. J. McGuire (2007), Earthquake swarms driven by aseismic creep in the Salton Trough, California, *J. Geophys. Res.*, **112**, 10.1029/2006JB004596.
- Maeda K. (1999), Time distribution of immediate foreshocks obtained by a stacking method, *Pure Appl. Geophys.*, **155**, 381–394.
- Marsan D., E. Prono, and A. Helmstetter, Monitoring aseismic forcing in fault zones using earthquake time series, submitted to *Bull. Seismol. Soc. Am.*, 2011.
- Melbourne T. I., and F. H. Webb (2002), Precursory transient slip during the 2001 $M_W = 8.4$ Peru earthquake sequence from continuous GPS, *Geophys. Res. Lett.*, **29**, 10.1029/2002GL015533.
- Miyazaki S., J. J. McGuire, and P. Segall (2011), Seismic and aseismic fault slip before and during the 2011 Tohoku earthquake, *Earth Planets Space*, in press.
- Mogi K. (1979), Two kinds of seismic gaps, *Pure Appl. Geophys.*, **117**, 1172–1186.
- Ogata Y. (1988), Statistical models for earthquake occurrences and residual analysis for point processes, *J. Am. Stat. Assoc.*, **83**, 9–27.
- Ogata, Y., and K. Katsura (1993), Analysis of temporal and spatial heterogeneity of magnitude frequency distribution inferred from earthquake catalogues, *Geophys. J. Int.*, **113**, 727–738.
- Ogata, Y., and K. Katsura (2006), Immediate and updated forecasting of aftershock hazard, *Geophys. Res. Lett.*, **33**, L10305, doi:10.1029/2006GL025888.
- Okada Y., K. Kasahara, S. Hori, K. Obara, S. Sekiguchi, H. Fujiwara and A. Yamamoto (2004), Recent progress of seismic observation networks in Japan - Hi-net, F-net, K-NET and KiK-net, *Earth Plan. Space*, **56**, xv–xxviii.
- Peng Z., J. E. Vidale, M. Ishii, and A. Helmstetter (2007), Seismicity rate immediately before and after main shock rupture from high-frequency waveforms in Japan, *J. Geophys. Res.*, **112**, 10.1029/2006JB004386.
- Reasenber P. (1999), Foreshock occurrence before large earthquakes, *J. Geophys. Res.*, **104**, 4755–4768.
- Sacks I. S., A. T. Linde, J. A. Snoke, and S. Suyehiro (1981), A slow earthquake sequence following the Izu-Oshima earthquake of 1978, in *Earthquake Prediction: An International Review*, eds. D. W. Simpson and P. G. Richards, Maurice Ewing series 4, Am. Geophys. Union.

Toda S., R. S. Stein, K. Richards-Dinger, and S. B. Bozkurt (2005), Forecasting the evolution of seismicity in southern California: Animations built on earthquake stress transfer, *J. Geophys. Res.*, **110**, 10.1029/2004JB003415.

Umino N., T. Okada, and A. Hasegawa (2002), Foreshock and aftershock sequence of the 1998 M 5.0 Sendai, northeast-

ern Japan, earthquake and its implications for earthquake nucleation, *Bull. Seismol. Soc. Am.*, **92**, 2465–2477.

Vidale J. E., and P. M. Shearer (2006), A survey of 71 earthquake bursts across southern California: Exploring the role of pore fluid pressure fluctuations and aseismic slip as drivers, *J. Geophys. Res.*, **111**, B05312, 10.1029/2005JB004034.

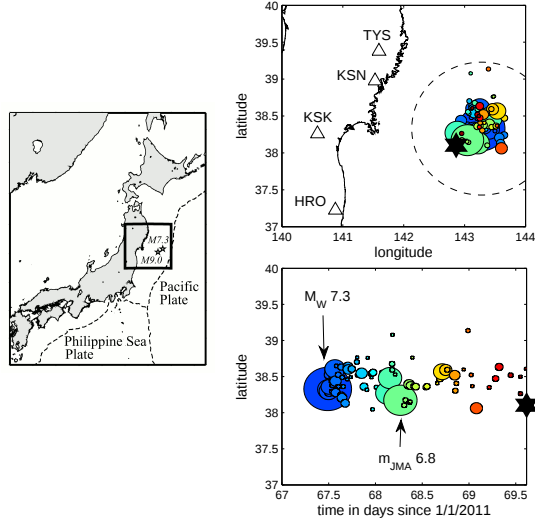


Figure 1. Setting of the study area (black rectangle) within Japan (left). The epicentral location (upper right) and relative occurrence time vs. latitude (bottom right) of all $m_{JMA} \geq 3.5$ earthquakes in the study area, from 3/9/2011 to the occurrence of the $M_W 9.0$ event (black star). The sizes of the circles are exponentially proportional to the JMA magnitudes. The color code highlights the temporal evolution of the sequence (from blue to red). The area within 100 km of the $M_W 7.3$ epicenter is outlined by dashed lines. Upper right plot shows the locations of the four F-net broadband stations used in this study.

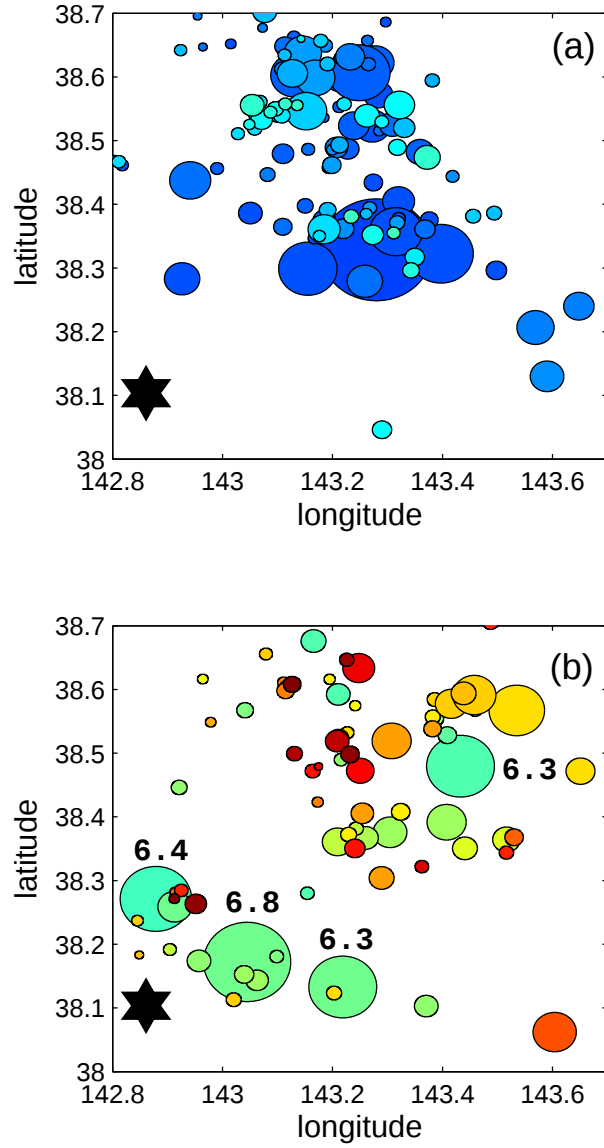


Figure 2. Epicenters of all JMA earthquakes occurring (top) from 3/9/2011 00:00 to 3/10/2011 03:00, and (bottom) from 3/10/2011 03:00 to the time of the $M_W 9.0$ earthquake (black star). The color code is the same as in Figure 1. Note the apparent migration of the seismicity towards the epicenter of the upcoming $M_W 9.0$ earthquake, and the cluster of three $m_{JMA} > 6$ events that delineates the maximum extent of this migration. The four large ($m_{JMA} > 6$) events occurring between 15 and 18 hours after the $M_W 7.3$ shock are indicated with their JMA magnitudes.

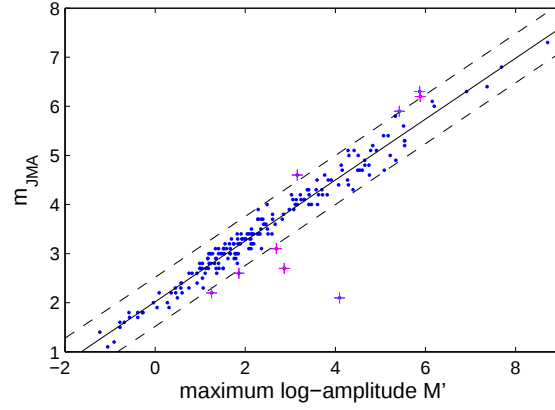


Figure 3. JMA magnitude m_{JMA} vs. maximum log-amplitude M' for the 200 earthquakes identified in both catalogs. The continuous black line shows the best linear fit, given by $m_{JMA} = 0.6212M' + 2.0204$. Dashed lines delimit the $m_{JMA} \pm 0.5$ interval around this linear trend. Only 9 events, shown with purple crosses, are more than 0.5 magnitude units away from the best linear fit.

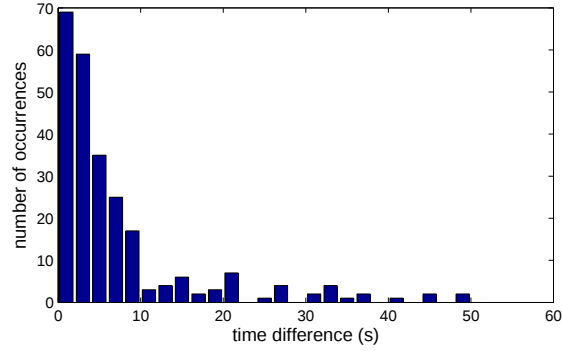


Figure 4. Distribution of the 245 time differences between the selected JMA earthquakes and the closest of the 1504 earthquakes identified by our algorithm.

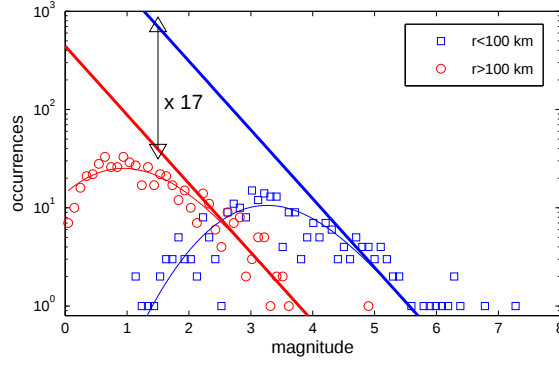


Figure 5. Number of occurrences vs. JMA magnitude for the two populations of earthquakes with distance r to the epicenter of the $M_W 7.3$ mainshock less or greater than 100 km. Thin lines: best fit with a Gutenberg-Richter law modulated by an error function (Ogata and Katsura, 1993, 2006; Daniel et al., 2008). The b -value is found equal to 0.7. Thick line: extrapolation of the Gutenberg-Richter law to low magnitudes. The number of earthquakes is found to be about 17 times greater for $r < 100$ km for all magnitudes.

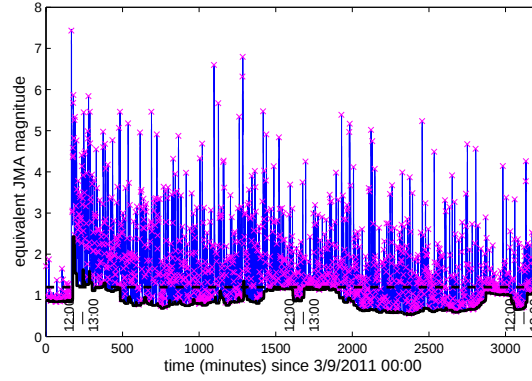


Figure 6. JMA magnitude-equivalent envelope $\mu(t)$, in blue, along with the 1504 detected earthquakes (purple crosses). The minimum noise level, shown with the thick black line, is computed by taking the minimum of $\mu(t)$ over a sliding window with 20 minute duration. Abrupt changes in this level are observed with a periodicity of 24 hours (see pattern around $t = 1500$ and $t = 3000$), and are likely due to human activities. The local 12:00 and 13:00 times are indicated with vertical lines. A magnitude cut-off at 1.2 (dashed black line) is thus imposed to avoid contamination by anthropogenic noise.

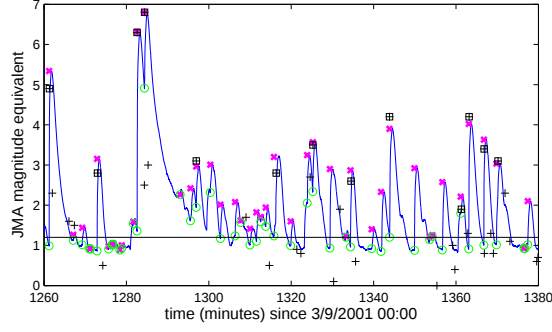


Figure 7. A 2 hour-long portion of the envelope $\mu(t)$ (blue), along with the JMA earthquakes (black crosses). The JMA earthquakes within 100 km of the epicenter of the $M_W 7.3$ shock are indicated with black squares. The earthquakes identified by our method are displayed at their occurrence times T , and magnitudes M (purple crosses). Their noise magnitude M_0 are shown with green circles. The cut-off magnitude at 1.2 is indicated by the horizontal line.

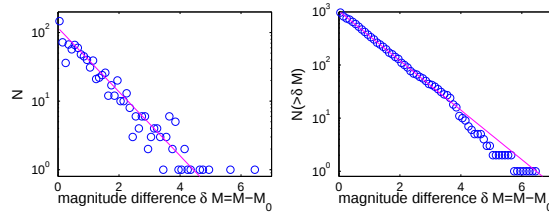


Figure 8. Number of occurrences vs magnitude difference $\delta M = M - M_0$ for the 979 earthquakes with $M > 1.2$. Left: number of events by magnitude intervals δM of width 0.1. Right: number of events greater than δM . In both graphs, the best fit giving $b = 0.47$ is shown in purple.

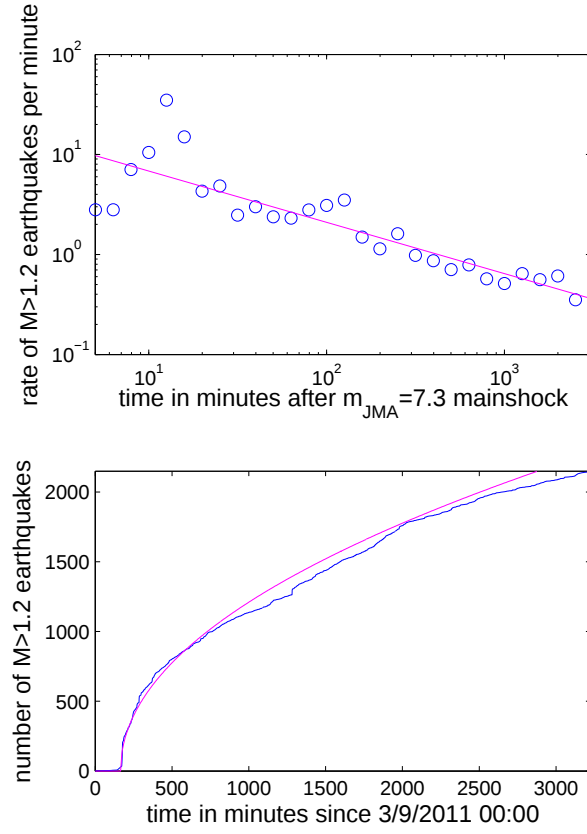


Figure 9. (Top) corrected rate of occurrence of $M > 1.2$ earthquakes after the $M_W 7.3$ mainshock. The first after-shock is detected 7 minutes after the mainshock, so that the first two plotted rate values (at 5 and 6.3 minutes) are identical. The best power-law fit is shown in purple, giving $\lambda(t) = 22.3 \times t^{-0.51}$. (Bottom) cumulative corrected number of $M > 1.2$ earthquakes, in blue. The integral of the best power-law fit for the top graph is shown in purple. A better overall fit is described in Section 4.

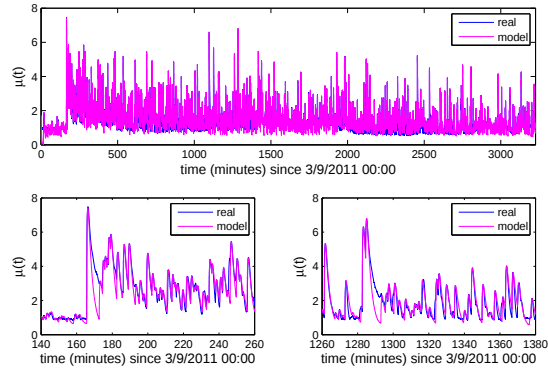


Figure 10. Comparison of the real and the modelled envelopes $\mu(t)$ for the real sequence. The two bottom graphs are zooms on 2 hour - long windows containing the $M_W 7.3$ (left) and the $m_{JMA} 6.8$ (right) shocks.

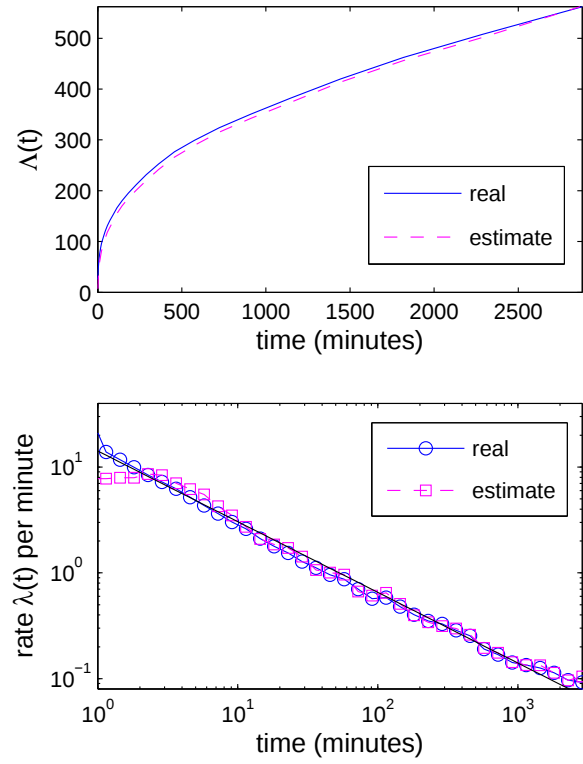


Figure 11. (Top) Poisson mean number $\Lambda(t)$ between the start of the sequence and time t , averaged over 200 synthetic realizations of the model. The true number (in blue) and the estimated number (purple dashed lines) are compared. (Bottom) same as for the top graph, but for the occurrence rate $\lambda(t)$. The black line is the best power-law fit in $t^{-0.67}$.

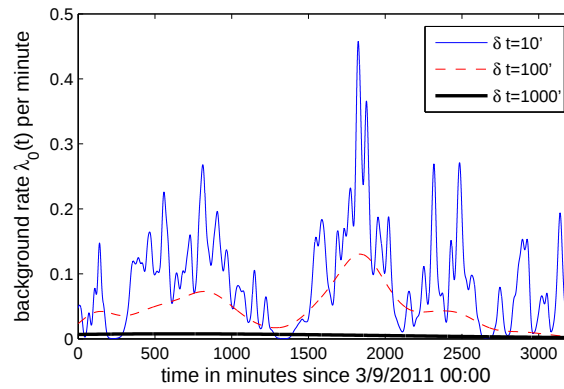


Figure 12. Background rate $\lambda_0(t)$ for three values of the smoothing parameter δt (in minutes).

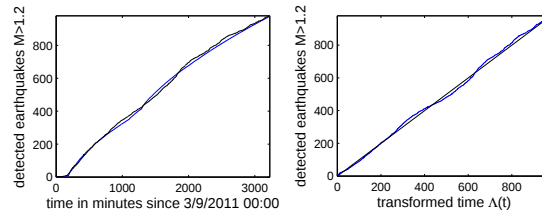


Figure 13. Best model (with $\delta t = 1000$ minutes) compared to the observed time series. (Left) number of detected earthquakes with $M > 1.2$ function of time, in black, and the modeled number, in blue. (Right) number of detected earthquakes with $M > 1.2$ function of transformed time $\Lambda(t) = \int_0^t dt \pi(t) \lambda(t)$ as given by the model, in blue. A good fit corresponds to this curve following the $y = x$ line, shown in black.

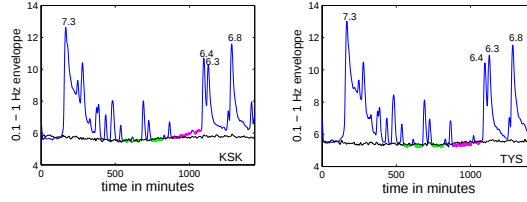


Figure 14. Plot of $s(t)$, the 1000s-smoothed logarithm of the envelope of the 0.1 - 1 Hz band-filtered vertical recording, for stations KSK (left) and TYS (right), for the 3/9/2011, in blue, and for the 3/8/2011, in black. The peaks correspond to large earthquakes, some of them labelled by their JMA magnitudes. Periods 1 and 2 are defined by the green and purple curves, respectively. The increase in noise is found for period 2 relative to period 1. It is particularly strong for station KSK, while much less significant for station TYS.

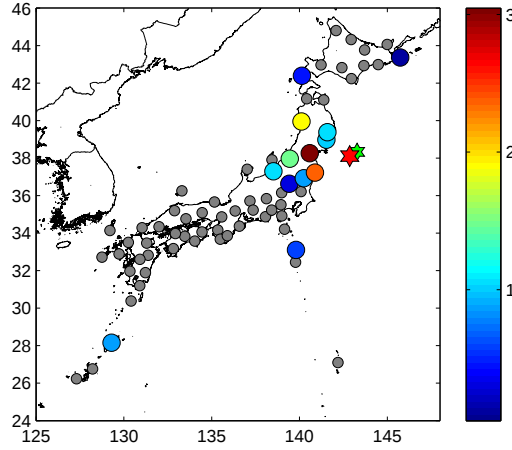


Figure 15. R -ratio measuring the increase in 0.1 - 1 Hz noise at all F-net stations. The stations with $R < 0$ are shown in grey, while those with $R > 0$ are colored according to the value of R . The two stations KSK and HRO are characterized by the strongest R ratio, greater than 3. The hypocenters of the $M_W 7.3$ and $M_W 9.0$ earthquakes are shown with the green and red stars, respectively.