Slow slip and aseismic deformation episodes associated with the subducting Pacific plate offshore Japan, revealed by changes in seismicity

D. Marsan[1], T. Reverso[1], A. Helmstetter[2], and B. Enescu[3]

D. Marsan, ISTerre, Université de Savoie, CNRS, Campus Scientifique, 73376 Le Bourget du Lac, France. (david.marsan@univ-savoie.fr)

T. Reverso, ISTerre, Université de Savoie, CNRS, Campus Scientifique, 73376 Le Bourget du Lac, France.

A. Helmstetter, ISTerre, CNRS, Université de Grenoble, France.

B. Enescu, University of Tsukuba, Earth Evolution Sciences, Faculty of Life and Environmental Sciences, Japan.

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

[2]ISTerre, CNRS, Université de Grenoble, France.

[3]University of Tsukuba, Earth Evolution Sciences, Faculty of Life and Environmental Sciences, Japan.

Abstract. Aseismic phenomena, including slow slip, can alter the surrounding seismicity. We here investigate how seismicity can be used in order to reveal episodes of aseismic deformation. An objective method is proposed, that accounts for both earthquake interactions and transient loading. Applying it to the 1990 - 2011 (pre-Tohoku) seismicity of the Japan subduction zone, we find several significant instances of aseismic transients. Small-scale and short duration transients are favored updip of the subducting plate. Large scale transients are mostly observed off-shore Ibaraki prefecture, in a partly decoupled zone that extends downdip. The four most intense of such transients have occurred periodically every 5.9 years, and are likely due to slow slip episodes. Other aseismic phenomena, including possible fluid intrusion in the outer-rise, are also detected. Finally, the seismicity in January and February 2011, close to the epicenter of the mega-thrust Tohoku earthquake, is found to be due to aseismic loading, confirming previous studies, although this transient is only one among others, and is not the most intense nor the most significant for the 21 year-long period studied here.

1. Introduction

Seismicity is an abundant observable which dynamics contains unique information on the stressing of faults, and how this stress evolves with time. Starting with Omori (1894), a rich corpus of observations and models have investigated the seismicity signature of sudden stress changes, more particularly as experienced following a major shock. More subtle and gradual changes in stress can also be estimated based on seismicity data alone (Marsan et al., 2013). This is the case of seismic swarms, either related to fluid intrusions (Hainzl and Ogata, 2005; Lombardi et al., 2006; Daniel et al., 2011) or to slow slip (Llenos and McGuire, 2011), for which the time evolution of the controlling deformation process is studied based on earthquake rates. To discriminate between 'normal' activity, i.e., earthquake occurrences due to both constant tectonic loading and stress steps imparted by previous earthquakes, from 'abnormal' activity, i.e., that includes episodic aseismic deformation, one must model the first ('normal' activity) and evaluate whether the residual activity is significant or not. If significant, it would indicate that the activation of extra seismicity is likely, revealing the presence of an underlying aseismic loading process.

In this paper, we develop an objective way of identifying aseismic transients by exploiting seismicity data. We define an aseismic transient as any episode of deformation with finite duration and extent that is not an earthquake. Only aseismic transients that cause seismicity activation are detectable with our treatment.

We here focus on earthquakes related to the Pacific plate subducting underneath Japan. Seismicity transients for the central part of this subduction, that ruptured with the 2011 $M_W 9.0$ Tohoku earthquake, are searched for. We limit ourselves to the 1990-2011 period for completeness purposes. Several significant seismicity transients are detected and characterized, including one related to activity in the two months prior to the 2011 mega-thrust earthquake.

2. Model

We model earthquake occurrences using the rate-density (number of earthquakes per unit time and unit area) $\lambda(x, y, t)$, defined as the sum of two contributions

$$\lambda = \mu + \nu \tag{1}$$

where ν accounts for earthquake interactions, i.e., triggering of earthquakes by previous earthquakes, and μ corresponds to the activity that would occur in the absence of any such interactions. This latter term is named the background rate-density, and is thought to model tectonic loading, as well as time-fluctuating forcing: fluid or magmatic intrusions, and slow slip events.

Earthquake interactions can be modeled by exploiting empirical and physical constraints on the triggering and nucleation of earthquakes. Many such models have been proposed in the past, sometimes with the aim of forecasting future seismicity rates.

Temporal dependence of triggering by an earthquake is modeled with the empirical Omori-Utsu law $\lambda_t(t) = \frac{K}{(t+c)^p}$. The cut-off time c is found to be of the order of minutes to tens of minutes when studying regional or world-wide seismicity, and can extend to several hours - tens of hours following strong shocks, as a result of incomplete detection.

Parameter K depends on mainshock magnitude. Numerous empirical analyses have shown that $K = K_0 e^{\alpha(m-m_0)}$, with typically $1.4 < \alpha < 2.5$ (Zhuang et al., 2004; Felzer et al., 2004; Helmstetter et al., 2005; Hainzl and Marsan, 2008). Lower values have been obtained for swarm seismicity (Hainzl and Ogata, 2005; Lombardi et al., 2006) but this possibly results from an under-estimation caused by inadequate modeling of the non-stationary loading during swarms (Marsan et al., 2013; Hainzl et al., 2013).

Spatial dependence is modeled using a power-law decay

$$\lambda_r(x,y) = \frac{\gamma-1}{2\pi} \times \frac{L^{\gamma-1}}{(r^2+L^2)^{\frac{\gamma+1}{2}}} \tag{2}$$

with $\gamma > 1$. This is in agreement with the distribution of hypocentral distances r between the mainshock and its direct aftershocks as studied by Marsan and Lengliné (2010): a power-law decay in $r^{-\gamma}$ for the linear density (Felzer and Brodsky, 2006) with $1.7 < \gamma < 2.1$ was obtained for r greater than the rupture length, although the challenging issue of only selecting direct aftershocks, and thus rejecting indirect aftershocks, generates significant uncertainties. This decay was shown (Marsan and Lengliné, 2010) to be consistent with static stress triggering governed by rate-and-state friction (Dieterich, 1994). The density of Equation (2) is normalized when summing over an infinite surface: $\int dr\ 2\pi r\ \lambda_r(r) = 1$, and decays at great r as $r^{-1-\gamma}$ which translates into a $r^{-\gamma}$ decay in linear density (equal to $2\pi r \lambda_r(r)$). We moreover model the influence length as $L = L_0 10^{0.5(m-m_0)}$, according to well established scaling laws (Wells and Coppersmith, 1994).

We finally integrate all these components into the modeled seismicity rate caused by earthquake interactions, expressed as:

$$\nu(x,y,t) = \frac{K_0(\gamma-1)}{2\pi} \sum_{i/t_i<t} \frac{e^{\alpha(m_i-m_0)}}{(t-t_i+c)^p} \times$$

$$\frac{\left[L_0\ 10^{0.5(m_i-m_0)}\right]^{\gamma-1}}{\left[(x-x_i)^2+(y-y_i)^2+L_0^2\ 10^{m_i-m_0}\right]^{\frac{\gamma+1}{2}}} \tag{3}$$

where the sum is performed on all earthquakes with index i occurring before time t. This is the simplest form that accounts for the observed dependence of triggering on distance, time, and mainshock magnitude. The linearity of the model, in particular, makes for a simple treatment; it amounts to separating and summing the triggering of the sources (seismic and aseismic). In this linear approach, the triggering caused by two mainshocks is the sum of the triggering caused by each one individually as if they were isolated.

It is important to emphasize that during periods of elevated aseismic loading, hence of high μ value, the interaction model is kept the same, i.e., the parameters γ, α, p, c, L_0 and K_0 are constant all throughout the duration of the dataset. Rigorously speaking, this assumption violates the prediction of the rate-and-state friction model, for which a time-varying background stress rate $\dot{\tau}$ induces a time-varying triggering kernel. However, the dependence of the triggering kernel on stressing rate, as predicted by the rate-and-state model, is difficult to investigate, and has never been convincingly demonstrated. Moreover, it can be shown that this model predicts that the total number of triggered aftershocks does not depend on stressing rate, in agreement with our model.

3. Data

We use the earthquake dataset of the Japan Meteorological Agency (JMA). The target region, for which we aim at detecting aseismic transients, correspond roughly to the rupture area of the 2011, $M_W9.0$ Tohoku mega-thrust earthquake: $36^o <$ latitude $< 41^o$, and $140.5^o <$ longitude $< 146^o$. We limit our study to the target period extending from 1/1/1990 to 9/3/2011, thus ending immediately before the occurrence of the $M_W7.3$ foreshock on the 2011 M_W9.0 earthquake. For this period and region, a $m_c = 3.5$ magnitude of completeness is obtained, which appears coherent with the completness maps of Nanjo et al. (2010). Seven earthquakes with $m \geq 7$ occurred in this area and period, the strongest shock being the 28/12/1994 M_W7.6 off Sanriku earthquake (Heki et al., 1997).

All earthquakes from 1/1/1970 to 9/3/2011 with $m \geq m_c$ and in an extended region (we add an extra 1^o in all directions) are kept as potential mainshocks, i.e., they are used as triggering sources and enter the summation of Equation (3). Including these events allows for a more accurate analysis (Wang et al., 2010), and avoids the spurious slowing down of background activity systematically observed when not doing so. Preliminary analyses have shown that adding these extra earthquakes, especially those occurring before the target period, is important for correctly estimating the background rate.

4. Searching for transient forcing

In section 2, we described how earthquake occurrences can be modeled as the sum of two terms μ and ν, the first accounting for aseismic loading, and the second being determined using a parameterized model of earthquake interactions. During a swarm of earthquakes, a large number of occurrences is observed, that cannot be explained solely by earthquake interactions. Such a sequence will thus temporary require a larger value of background rate-density μ. On the contrary, an aftershock sequence can be explained by interactions between earthquakes, and should therefore be characterized by a normal μ value, i.e., no temporary increase of μ is required to explain the data. The rate-density μ and its temporal fluctuations thus has the power to reveal aseismic loading transients. Given our aim at finding episodes of anomalously high aseismic loading that potentially correspond to slow slip events, we therefore need to search for locally high μ values, that are significantly higher than normal. The meaning of 'normal' and 'significantly higher' are defined in this section.

We define two families of models: model 0 is the null hypothesis of 'no transient in loading rate'. It models seismicity with equations (1) and (3), with $\mu(x, y, t) = \mu_0(x, y)$, hence a constant background rate density at all locations.

Model 1 is a modification of model 0, obtained by adding a temporary and local change in loading μ. It uses the same parameterization for the interaction term ν, and models $\mu(x, y, t)$ as $\mu_0(x, y)$ almost everywhere except for a circular zone of radius L centered on a location (x^*, y^*) and for a period $t^* - \tau/2 < t < t^* + \tau/2$:

$$\mu(x, y, t) = \mu^* \text{ if } (x, y, t) \in D_{L,\tau}(x^*, y^*, t^*) \quad (4)$$

$$\mu(x, y, t) = \mu_0(x, y) \; otherwise \quad (5)$$

where $D_{L,\tau}(x^*, y^*, t^*)$ stands for the circular area with radius L and time interval of duration τ centered on (x^*, y^*, t^*).

In order to detect significant loading transients, we proceed as follows:

- Step 1: We optimize the parameters $\{\alpha, K, p, c, L_0, \gamma\}$ entering the formulation of ν, see Equation (3), given the data, and also optimize $\mu_0(x, y)$.

- Step 2: We define a model of type 1 with a transient loading at $D_{L,\tau}(x_i, y_i, t_i)$ centered on the location and time of earthquake i. Given ν and μ_0 of Step 1, we optimize a model of type 1 independently for each earthquake. We here emphasize that we therefore investigate the possibility of as many loading transients as there are earthquakes in the dataset.

• Step 3: We compare model 0 to all models 1. To do so, we test whether a model 1 is significantly better than model 0 to explain the data.

We now detail each of these three steps. In step 1, we search for the maximum likelihood estimate (MLE) of the parameters entering the formulation of ν. Because the value of c has a weak impact on the results as long as it remains within a realistic interval, we simplify the optimization by setting $c = 0.001$ day ($\simeq 1.5$ minutes). We also optimize $\mu_0(x, y)$, using the method of Zhuang et al. (2002): (i) an a priori guess of $\mu_0(x, y)$ is made; (ii) parameters $\{\alpha, K, p, L_0, \gamma\}$ are optimized, given this μ_0; (iii) the background probabilities $\omega_i = \frac{\mu_0(x_i,y_i)}{\mu_0(x_i,y_i)+\nu(x_i,y_i)}$ are computed for each earthquake i, using the MLE parameters of (ii) for computing ν; (iv) the a posteriori $\mu_0(x, y)$ is obtained by smoothing the probabilities ω_i with the smoothing scale ℓ:

$$\mu_0(x, y) \;=\; \sum_i \omega_i \; e^{-\sqrt{(x-x_i)^2+(y-y_i)^2}/\ell} / 2\pi\ell^2 T \qquad (6)$$

where T is the total duration of the earthquake catalogue (7737 days from 1/1/1990 to 9/3/2011). This a posteriori is then taken as a new a priori, and steps (ii) to (iv) are run again iteratively until convergence of all inverted parameters. The solution does not depend on the initial a priori guess of μ_0, as long as it is non-zero.

Table 1 lists all parameter values inverted for $\ell = 20$ km and $\ell = 80$ km. The α values are small in both cases. This could be caused by the anisotropy of aftershocks spatial distribution (Hainzl et al., 2008) or by the dominant contribution of aseismic episodes (Hainzl et al., 2013). We however think that, at the scale of this study, i.e., 21 years and $\sim 500 \times 500$ km^2, aseismic transients do not dominate the seismicity, as confirmed in Section 5. For comparison, in their analysis of Japanese seismicity at larger scale

($121^o - 155^o$ of longitude, $21^o - 48^o$ of latitude, 1/1/1926 - 31/12/1999, $m \geq 4.2$), Zhuang et al. (2004) found $\alpha = 1.33$ to 1.36 depending on the choice of the spatial kernel. As a consequence of a low α value, the model has greater flexibility to account for changes in background rate. Our estimation of aseismic transients is therefore conservative: using a greater, more typical α value would result in a greater number of detected transients, with greater significance. We discuss in Section 5 the dependence of our results on the estimate of α.

The background rate density $\mu_0(x, y)$ is shown for $\ell = 20$ km in Figure 1; the one obtained with $\ell = 80$ km is only a smoother version of it. The distribution of μ_0 is not trivial: the largest values are found on the subduction interface in the south-westernmost corner of the studied area, although the strongest activity is located more to the north at about latitude 39^o to 40^o. This strong activity is mostly due to the aftershocks of the 1994 M_W7.6 off Sanriku earthquake, and therefore does not contribute to the background activity. We do not search for an optimal smoothing length ℓ as in Zhuang et al. (2002) since this length will be ultimately linked to the size of the loading transients we aim to detect.

In step 2, we only need to estimate μ^*, the background rate density of the region - period $D_{L,\tau}(x^*, y^*, t^*)$. This is done independently for each earthquake, by centering $D_{L,\tau}$ on one earthquake at a time. The MLE of μ^* is found numerically by minimizing $J(\mu^*) = \mu^* \pi L^2 \tau - \sum_j \ln(\mu^* + \nu_j)$, where j are the indices of all earthquakes in $D_{L,\tau}(x^*, y^*, t^*)$. For coherence we fix L so that $D_{L,\tau}$ has the same surface πL^2 as the surface integral ($2\pi\ell^2$) of the ponderation $e^{-r/\ell}$ used in Equation (6). This imposes that $L = \ell\sqrt{2}$. Choosing L independently of ℓ is possible, but this would add an extra parameter; moreover, as only transients with size of L or greater can be detected here, choosing $\ell > L$ would imply too strong a smoothing when computing μ_0 in step 1. Parameter τ (duration of the transient) is a free parameter chosen by the observer.

Because any model 1 is more flexible than model 0, it always fits the data better. The significance of the improvement is measured by computing the change in Akaike Information Criterion ΔAIC (Akaike, 1974), here defined as

$$\Delta\mathrm{A}IC \;=\; 2\;(J_1 - J_0 + 5) \tag{7}$$

where J_i is the minimum of the cost function for model i, and a penalty of 5 is applied because 5 extra parameters are needed in models 1: $x^*, y^*, t^*, \mu^*, \tau$. A negative ΔAIC is required for the loading transient to be significant compared to the null hypothesis of no change in loading rate.

We finally merge together transients that are significant and that overlap (at least one earthquake in common). We end up with zero or several transients of sizes $\geq \ell$ and duration $\geq \tau$, that do not overlap. The corresponding model has $\mu(x, y, t) = \mu_0(x, y)$ outside these transients and $\mu(x, y, t) = \mu_i(x, y, t)$ if $\{x, y, t\}$ is within transient number i, where μ_i is the optimized background rate density for this transient. This model is the model with the lowest possible ΔAIC constructed by merging disks $D_{L,\tau}$ centered on earthquakes.

5. Aseismic transients

The resulting $\mu(x, y, t)$ depends on scale parameters ℓ and τ. To illustrate this dependence, and explore some of the spectrum of transient sizes and durations, we describe the results obtained with $\ell = 20$ km, $\tau = 40$ days, and with $\ell = 80$ km and $\tau = 100$ days. We characterize the intensity of a transient by the ratio $\rho = \frac{\mu(x,y,t)}{\mu_0(x,y)}$, i.e., how much the local background rate density is increased during the transient. While $\rho < 1$ is a possibility, i.e., a transient shutdown of the background activity, it is not observed here at both choices of ℓ and τ.

5.1. Small-scale transients ($\ell = 20$ km, $\tau = 40$ days)

We find 39 such distinct transients, with ρ ranging from 14 to 676. Figure 2 displays their distribution in time, and Figure 3 in space. Apart from a 4 year-gap (1999-2003), their occurrences span the whole period, with a clear overall slowing down, see Figure 4.

Anomalous aftershock sequences: Among these transients, 6 are clearly related to aftershock sequences (magenta crosses in Figure 3). Although the method is designed to account for earthquake interactions, and therefore to model aftershock sequences without requiring an increase in loading rate, it only does so on the basis of triggering kernels that are mean laws, cf. section 2. Such an approach works well when summing over many mainshocks, but fails to account for variability around this mean behavior. This is here the case following larger mainshocks, and, more significantly, when focusing at small scale. As will be explained below, larger scale transients do not include any aftershock sequences. Moreover, small-scale transients related to aftershocks discriminate sequences that are remarkably vigorous in terms of activity, again as compared to the mean behavior as expressed by $K = K_0 e^{\alpha(m-m_0)}$. This is possibly linked to high post-seismic slip rates.

On a side note, the strong transient $\rho \simeq 680$ at 2005.9 corresponds to the aftershock sequence of the 2005/11/14 $M_W 7.0$ ($m_{JMA} 7.2$) earthquake located in the outer-rise at

$\simeq 38^o$ latitude and $\simeq 145^o$ longitude. Quite remarkably, no aftershock is observed in the immediate vicinity of this mainshock; the compact cluster of aftershocks is about 30 to 40 km away to the north-west. Error in localization of the mainshock is unlikely as other catalogs (USGS PDE, ISC, and Harvard CMT) are in agreement with the epicenter provided by JMA. This large distance between the mainshock and the aftershock cluster is anomalous for the model, which therefore requires to significantly increase the local loading rate to explain this cluster.

We discard transients related to aftershock sequences by requiring that the maximum magnitude of all earthquakes in a transient $D_{L,\tau}$ is less than 6. This simple criterion is here sufficient, see Figure 3, bottom left graph.

Outer-rise transients: Episodes of high aseismic forcing mostly affect two distinct zones: (1) at the plate boundary between the subducting and overriding plates, between about 37.5^o and 38.5^o, delimiting a band parallel to the trench, at 40 km to 80 km from it (about 15 km to 30 km downdip on the subduction interface); and (2) in the outer-rise, at 20 km to 60 km from the trench. The extensional regime in the outer-rise can allow for the episodic intrusion of lower crustal fluids, or alternatively rapid pore fluid pressure change could be related to penetrating seawater as suggested for other outer-rise swarm-like seismic activity (Tilmann et al., 2008), hence a very different mechanism than aseismic slip that better explains the observed clusters on the subduction interface. Among the non-aftershock sequence transients, the strongest one, labelled 5 in Figure 3, is particularly intense, with $\rho \simeq 650$. It occurs in 1995, and is well isolated. Its characteristics are given in more details in Figure 5. The maximum magnitude is 4.6, which cannot explain the abrupt increase in rate at the time of the transient, especially as it is the 10th to occur in this swarm; 32 earthquakes occurred in this zone in 40 days, twice as many as during the rest of the 21 year interval (47 $m \geq 3.5$ earthquakes in total, for 21 years, including the transient). No clear migration of seismicity is observed during this transient. We moreover note that the mean depth of these earthquakes is 40.3 km, i.e., so that this transient is unlikely to be caused by penetrating seawater (Tilmann et al., 2008).

Activity prior to the 2011, $M_W 9.0$ Tohoku earthquake: A relatively high level of seismicity lasted for about one month, between mid-january to mid-february 2011, in a zone of about 40 km length, close to the epicenter of the impending $M_W 9.0$ Tohoku earthquake. This activity was reported by Ando et al. (2011) and Kato et al. (2012) as a precursory phenomenon related to the mainshock, mainly owing to its proximity in time and space, and to the apparent migration of the earthquakes towards the epicenter of the mainshock. It was suggested by these authors that such an activity could be due to ongoing slow slip on the subduction interface, that further loaded the asperity which failure was to initiate the mainshock. In contrast, the 2.5 day long foreshock sequence generated by the strongest foreshock, a $M_W 7.3$ shock, does not seem to be characterized by anomalous forcing, although the signature of slow slip in the seismicity could be hidden in the high aftershock activity following the $M_W 7.3$ earthquake (Marsan and Enescu, 2012).

Among the 39 transients, the last one is effectively related to the precursory activity of January - February 2011. It contains the five $m \geq 3.5$ earthquakes occurring within 20 days (i.e., $\tau/2$) and 20 km of the $m = 3.5$ 19/1/2011 earthquake, at longitude 143.17 and latitude 38.19, see Figure 6, and has an intensity of $\rho = 64$, i.e., the loading rate must be 64 times that of the long-term average rate. This corresponds to the initial phase of the precursory activity, as it excludes the mid- to end-february activity (9 earthquakes with $m \geq 4.0$ in the same zone).

We test whether our results are robust regarding to the low α value obtained when optimizing the model parameters. To do so, we re-run the analysis, imposing an $\alpha = 2$ value. The small scale transients are effectively little affected by this change. More precisely, (1) the six anomalous aftershock clusters are again detected; (2) the outer-rise transient is still found with a great significance; (3) the non-aftershock related transients are in the updip part of the subduction zone, and within $\simeq 30$ km of the trench on the outer-rise side; (4) the pre-Tohoku transient is again detected, with a significance level that classify it as the 7th most significant, non-aftershock transient, as compared to being the 6th with $\alpha = 1.05$.

5.2. Large-scale transients ($\ell = 80$ km, $\tau = 100$ days)

19 transients are observed at large scale, see Figures 7 and 8. Unlike transients at small scale, none appears related to well-identified aftershock sequences. Their spatial distribution is quite different from those at small scale, with a clear tendency to cluster on the active subduction interface in the south-westernmost corner of the studied area, which coincides with the zone with the highest background rate density μ_0, see Figure 1. According to our analysis, this zone was, during the 1990-2011 period, the most seismically decoupled, as testified by Figure 1. This is in agreement with the study by Uchida and Matsuzawa (2011), who used repeating earthquakes to infer that the northern part is strongly coupled ($\simeq 100\%$) while the southern part is partly decoupled ($\simeq 70\%$ according to Figure 2 of Uchida and Matsuzawa, 2011). It is also the same zone that underwent the most significant large-scale increases in loading. This shows that the seismicity loading

rate on a partly decoupled subduction zone fluctuates significantly with time, and that these fluctuations affect large portions of the interface, i.e., they are not localized at small scale.

We now focus on the most intense transients at $\ell = 80$ km, by keeping the 8 that have $\rho > 35$ (this threshold is here arbitrary), and on the 15 most intense transients at $\ell = 20$ km (the 9 highlighted in Figure 3, plus the 6 related to aftershock sequences). We find very little overlap between the two scales: part from one common transient in 1998, none of these 8 and 15 transients have common earthquakes. This indicates that the processes triggering episodic loading transients at these two scales likely have distinct natures. The small-scale transients have stronger intensities, and are found in aftershock sequences, and on both sides of the trench and close to it, so that the deeper part of the subducting plate is devoid of them. If slow slip is responsible for transients on the subduction interface, then only downdip transients can develop to reach large spatial extents, and longer durations.

Out of the 8 most intense large-scale transients, six are located in the south, while only two (number 2 and 3) are in the north, see Figure 8. The 6 southern transients are partly clustered, with two pairs (4 and 5; 6 and 7) occurring very closely in space and time. These two pairs would have merged into single transients had the two scales ℓ or τ been larger. Considering these two pairs as single transients, we note that the occurrence times of the remaining four intense, large scale transients in the southern part of the subduction zone are very regular, with a cycle of 5.9 years (Figure 9). We have no precise model nor explanation for such a periodic behavior. This period is within the range of recurrence times observed for the Boso slow slip event (Hirose et al., 2012), although the recurrence is here much more periodic. We also note that this feature is not robust if imposing $\alpha = 2$, a more standard value for regional seismicity, while the others are little affected by this change of α. Indeed, when taking $\alpha = 2$, the 8 most significant transients are now all in the south, and the regular pattern of large-scale transient occurrences in the south is perturbed by the addition of the two new transients, that are the least significant of this group.

We show in Figure 10 the most intense large scale transient, which affects the partly decoupled zone. It encompasses 68 $m \geq 3.5$ earthquakes, between the 1998/6/29 and 1998/12/22. The raw rate of earthquakes is 143 $m \geq 3.5$ per year during this period, an 60% increase from the 1990 - 2008 average (the seismicity rate is then perturbed in 2008 by a magnitude 7 earthquake). The maximum magnitude is 5.3 during the transient, so that aftershock activity alone cannot explain this increase. We furthermore find that this increase in background rate is much stronger, by a factor of $\rho = 56$. This increase in background rate density explains 25 of the 68 earthquakes, the others being aftershocks of previous events according to the model, while only 0.44 background earthquakes would be normally expected for this zone and this duration, if the background loading was indeed constant.

6. Discussion and Conclusions

Our limited capacity to detect slow slip events is particularly highlighted by the existence of tremor episodes without resolved transient surface displacements in regions where episodic tremor and slip (ETS) events are known to occur (e.g., Kao et al., 2009). The possibility of using seismicity to reveal aseismic slip is therefore appealing. However, the relationship between increased slip and changes in seismicity is not straightforward, as already suggested by Pollitz and Johnston (2006). For example, ETS events in Cascadia have long been recognized as uncorrelated with detectable changes in seismicity (Dragert et al., 2001; Schwartz and Rokosky, 2007), although a more recent and focused study by Vidale et al. (2011) points to a possible, albeit weak, activation following a moderate tremor episode.

Two slow slip episodes have been documented for the 2008 - 2011 (pre-Tohoku) period using pressure gauges installed offshore at latitude $\simeq 38.2^o$ to 38.4^o (Ito et al., 2012). The first occurred in November 2008, and lasted for a week. We note that the removal of the instrumental drift on the pressure gauge measurements in late November 2008 (around julian day 330) rather than at the beginning of the suspected deformation episode at julian day 320 (Figure 4 of Ito et al., 2012) could have created an artificial pressure transient. However, keeping with the hypothesis that an actual transient effectively took place in November 2008, no significant direct seismicity activation is observed during this period, although a magnitude 6.1 earthquake occurred 10 days later. It is not clear whether this earthquake has anything to do with the suspected slow slip. This slow episode is not detected with our method, simply because the $m = 6.1$ earthquake and its aftershock sequence are classified as 'normal' by our model, which does not require any substantial increase in loading rate to explain them. The second episode of slow slip found by Ito et al. (2012) corresponds to the activity in January and February 2011, prior to the $M_W 9.0$ mega-thrust earthquake. This event is well detected by our method, see Section 5.1, as anomalous extra seismicity is generated by it. This transient is however not the most significant in the 1990-2011 period, showing that the use of aseismic transient detection methods to anticipate the occurrence of strong or giant earthquakes is not straightforward.

The method proposed here cannot therefore exhaustively detect all slow slip episodes, as some can occur without any significant seismicity changes. Moreover, a significant increase in background activity does not necessarily imply slow slip: fluid intrusions, as expected with the 1995 outer-rise transient of Figure 5, can also trigger such anomalous activity, even in the context of subduction zones.

Small transients have been found to occur preferentially in the upper part of the subducting plate, while the large transients cover also the downdip portion. This is in agreement with models that describe the frictional properties of the Japan subduction zone as a mixture of velocity-weakening asperities and conditionnally-stable sliding patches (Schwartz and Rokosky, 2007). The density of creeping patches increasing with depth (Uchida and Matsuzawa, 2011), it is easier for slow slip to develop to larger extent downdip rather than updip. Moreover, this larger extent also implies longer durations, as slow slip can migrate or diffuse over a larger area. We therefore think that our observation that small-scale and large-scale transients are mostly distinct in their spatial and temporal distributions is a consequence of the depth-dependent frictional properties.

Past observations suggest that there exists a continuum of slip modes, from rapid (seismic) to slow (aseismic), with slip events spanning some of this continuum as they evolve (Peng and Gomberg, 2010). When searching for slow slip events, it is therefore expected that the result must depend on the scale of observation: a time-fluctuating slip or loading rate will exhibit different transients at different scales. A limit of the method proposed here is that the observation scale must be imposed a priori, while a more sophisticated algorithm could locally optimize this scale to reveal at once this continuum of scales. Such a development will be the aim of future methodological work.

Acknowledgments

This work was supported by the French ANR ASEISMIC project and the ECFP7 REAKT project. We would like to thank Sebastian Hainzl for discussions and help on this topic. We used the earthquake catalog arranged by the Japan Meteorological Agency in cooperation with the Ministry of Education, Culture, Sports, Science and Technology.

Bibliography

Akaike, H. (1974), A new look at the statistical model identification, *IEEE Trans. Aut. Cont.*, **19**, 716 – 723.

Ando R., and K. Imanishi (2011), Possibility of M_W9.0 mainshock triggered by diffusional propagation of afterslip from M_W7.3 foreshock, *Earth Plan. Space*, **63**, 767.

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 20032004 Ubaye seismic swarm, France, *J. Geophys. Res.*, **116**, B01309, doi: 10.1029/ 2010JB007551.

Dieterich, J. (1994), A constitutive law for rate of earthquake production and its application to earthquake clustering, *J. Geophys. Res.*, **99**, 2601 – 2618.

Dragert H., K. Wang, and T. S. James (2001), A slient slip event on the deeper Cascadia subduction interface, *Science*, **292**, 1525–1528.

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.

Felzer, K. R., and E. E. Brodsky (2006), Decay of aftershock density with distance indicates triggering by dynamic stress, *Nature*, **441**, 735 – 738.

Hainzl, S., and Y. Ogata (2005), Detecting fluid signals in seismicity data through statistical earthquake modeling, *J. Geophys. Res.*, **110**, B05S07, doi: 10.1029/2004JB003247

Hainzl, S., and D. Marsan (2008), Dependence of the Omori-Utsu law parameters on main shock magnitude: Observations and modeling, *J. Geophys. Res.*, **113**, B10309, doi:10.1029/2007JB005492.

Hainzl, S., A. Christophersen, and B. Enescu (2008), Impact of Earthquake Rupture Extensions on Parameter Estimations of Point-Process Models, *Bull. Seismol. Soc. Am.*, **98**, 2066 – 2066, doi: 10.1785/0120070256.

Hainzl, S., O. Zakharova, and D. Marsan (2013), Impact of Aseismic Transients on the Estimation of Aftershock Productivity Parameters, *Bull. Seismol. Soc. Am.*, **103**, doi: 10.1785/0120120247.

Hayes G. P., D. J. wald, and R. L. Johnson (2012), Slab1.0: A three-dimensional model of global subduction zone geometries, *J. Geophys. Res.*, **117**, doi: 10.1029/2011JB008524.

Heki K., S. Miyazaki, and H. Tsuji (1997), Silent fault slip following an interplate thrust earthquake at the Japan Trench, *Nature*, **386**, 595 – 598.

Helmstetter, A., Y. Y. Kagan, and D. D. Jackson (2005), Importance of small earthquakes for stress transfers and earthquake triggering, *J. Geophys. Res.*, **110**, B05S08, doi:10.1029/2004JB003286.

Hirose, H., H. Kimura, B. Enescu, and S. Aoi (2012), Recurrent slow slip event likely hastened by the 2011 Tohoku earthquake, *Proc. Nat. Acad. USA*, **109**, 15157 – 15161.

Ito Y., R. Hino, M. Kido, H. Fujimoto, Y. Osada, D. Inazu, Y. Ohta, T. Iinuma, M. Ohzono, S. Miura, M. Mishina, K. Suzuki, T. Tsuji, and J. Ashi (2012), Episodic slow slip events in the Japan subduction zone before the 2011 Tohoku-Oki earthquake, *Tectonophysics*, http://dx.doi.org/10.1016/j.tecto.2012.08.022

Kao H., S. J. Shan, H. Dragert, and G. Rogers (2009), Northern Cascadia episodic tremor and slip: A decade of tremor observations from 1997 to 2007, *J. Geophys. Res.*, **114**, B00A12.

Kato A., K. Obara, T. Igarashi, H. Tsuruoka, S. Nakagawa, and N. Hirata (2012), Propagation of slow slip leading up to the 2011 M_W9.0 Tohoku-Oki earthquake, *Science*, **335**, 705 – 708.

Llenos, A. L., and J. J. McGuire (2011), Detecting aseismic strain transients from seismicity data, *J. Geophys. Res.*, **116**, B06305, doi: 10.1029/ 2010JB007537.

Lombardi, A.,W. Marzocchi, and J. Selva (2006), Exploring the evolution of a volcanic seismic swarm: The case of the 2000 Izu Islands swarm, *Geophys. Res. Lett.*, **33**, L07310, doi: 10.1029/2005GL025157.

Marsan, D., and O. Lengliné (2010), A new estimation of the decay of aftershock density with distance to the mainshock, *J. Geophys. Res.*, **115**, B09302, doi: 10.1029/2009JB007119.

Marsan D., and B. Enescu (2012), Modeling the foreshock sequence prior to the 2011, M_W9.0 Tohoku, Japan, earthquake, *J. Geophys. Res.*, **117**, B06316, doi:10.1029/2011JB009039.

Marsan D., E. Prono, and A. Helmstetter (2013), Monitoring Aseismic Forcing in Fault Zones Using Earthquake Time Series, *Bull. Seismol. Soc. Am.*, **103**, 169 – 179, doi: 10.1785/0120110304.

Nanjo K. Z., T. Ishibe, H. Tsuruoka, D. Schorlemmer, Y. Ishigaki, and N. Hirata (2010), Analysis of the completeness magnitude and seismic network coverage of Japan, *Bull. Seismol. Soc. Am.*, **100**, 3261 – 3268.

Omori, F. (1894), On aftershocks, *Rep. Imp. Earthquake Invest. Comm.*, **2**, 103 – 109.

Peng Z., and J. Gomberg (2010), An integrated perspective of the continuum between earthquakes and slow-slip phenomena, *Nature Geosc.*, **3**, 599 – 607.

Pollitz F. F., and M. J. S. Johnston (2006), Direct test of static stress versus dynamic stress triggering of aftershocks, *Geophys. Res. Lett.*, **33**, L15318, doi:10.1029/2006GL026764.

Schwartz S. Y., and J. M. Rokosky (2007), Slow slip events and seismic tremor at circum-Pacific subduction zones, *Rev. Geophysics*, **45**, RG3004.

Tilmann F. J., I. Grevemeyer, E. R. Flueh, T. Dahm, and J. Gossler (2008), Seismicity in the outer rise offshore southern Chile: Indication of fluid effects in crust and mantle, *Earth Plan. Sc. Lett.*, **269**, 41 – 55.

Uchida N., and T. Matsuzawa (2011), Coupling coefficient, hierarchical structure, and earthquake cycle for the source area of the 2011 off the Pacific coast of Tohoku earthquake inferred from small repeating earthquake data, *Earth Plan. Space*, **63**, 675 – 679.

Vidale J. E., A. J. Hotovec, A. Ghosh, K. C. Creager, and J. Gomberg (2011), tiny intraplate earthquakes triggered by nearby episodic tremor and slip in Cascadia, *Geoch. Geophys. Geosyst.*, **12**, Q06005.

Wang Q., D. D. Jackson, and J. C. Zhuang (2010), Missing links in earthquake clustering models, *Geophys. Res. Lett.*, **37**, L21307.

Wells D. L., and K.J Coppersmith (1994), New emperical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement, *Bull. Seismol. Soc. Am.*, **84**, 974 – 1002.

Zhuang, J., Y. Ogata, and D. Vere-Jones (2002), Stochastic declustering of space-time earthquake occurrences, *J. Am. Stat. Assoc.*, **97**, 369 – 380.

Zhuang J., Y. Ogata, and D. Vere-Jones (2004), Analyzing earthquake clustering features by using stochastic reconstruction, *J. Geophys. Res.*, **109**, B05301.

ℓ	α	p	L_0	γ	K_0	$\bar{\omega}$
20 km	1.05	0.93	0.81 km	2.47	0.019	0.22
80 km	0.98	0.89	0.91 km	2.31	0.029	0.068

Table 1. Optimized ETAS parameters for the null hypothesis (model 0) at two distinct smoothing scales ℓ. The mean probability $\bar{\omega}$ of being a background earthquake is not a parameter per see, and is only given here for information. Parameter L_0 is for $m_0 = 3.5$.

Figure 1. Background rate density for $\ell = 20$ km, in number of $m \geq 3.5$ earthquakes per year per 100×100 km^2. For clarity, we only display earthquakes with $m \geq 4.5$ (white dots), although all $m \geq 3.5$ earthquakes were effectively used in our computations.

Figure 2. Temporal distribution of the 39 loading transients found with $\ell = 20$ km and $\tau = 40$ days (blue dots). The thick red dots and labels mark the 9 most intense transients not related to aftershock sequence ($\rho > 50$), while magenta crosses are for transients related to aftershock sequences ($m_{max} > 6$). The maximum magnitude m_{max} is among all earthquakes in the transient.

Figure 3. Map of all $m \geq 3.5$ earthquakes, 1/1/1990 - 9/3/2011. The transients are depicted using the same symbols and labels as with Figure 2. The anomalous aftershock sequence of 2005 is located between transients 2 and 8. Magnitude $m \geq 6$ earthquakes are shown with circles which radii equal $L_0 \times 10^{0.5(m-3.5)}$. The trench (Hayes et al., 2012) is shown with the thin black line.

Figure 4. Occurrence times of the 39 loading transients found with $\ell = 20$ km and $\tau = 40$ days.

Figure 5. Characteristics of the 1995 outer rise transient, labelled 5 in Figure 3. The inset in map shows a zoom in on all earthquakes in the zone, those occurring during the transient colored in red. The circles have diameters equal to the rupture length $2L_0 \times 10^{0.5(m-3.5)}$ (in km). The spatial distribution is too diffuse to be explained by earthquake interactions. Right graphs: number and magnitude of all earthquakes in the rectangular zone of the inset, still with those during the transient colored in red.

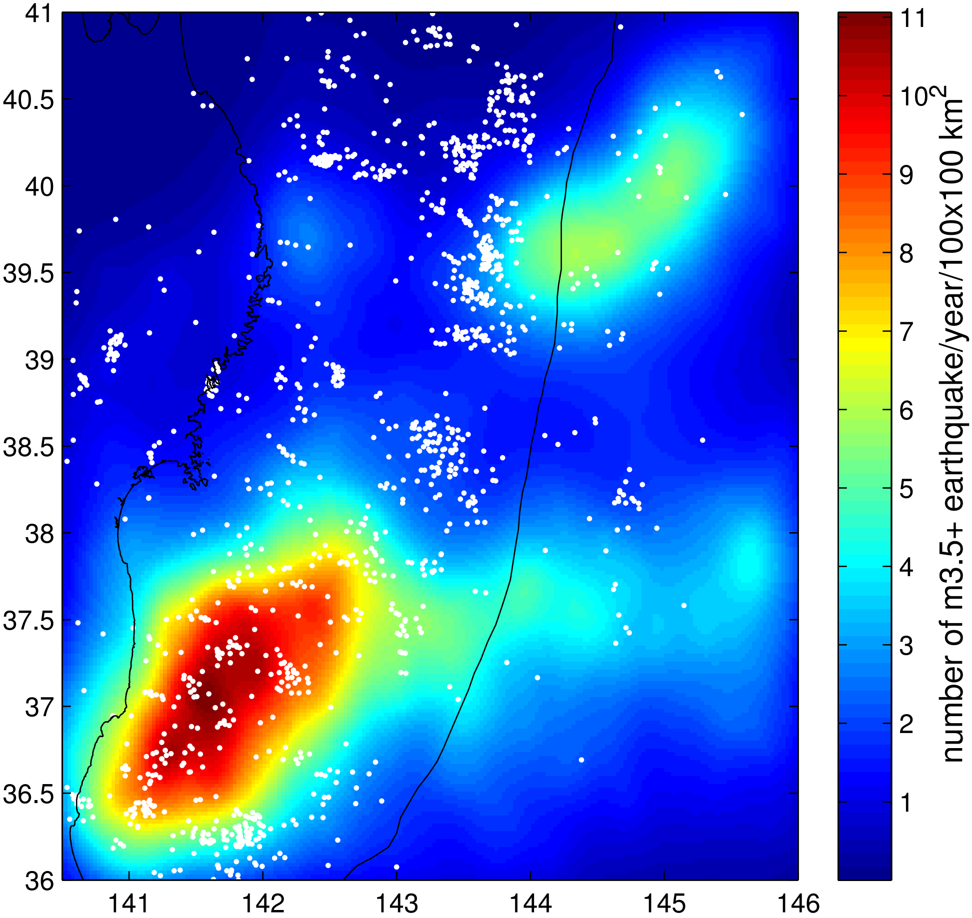
Figure 6. Same as Figure 5, but for the pre-Tohoku transient, labelled as 9 in Figure 3.

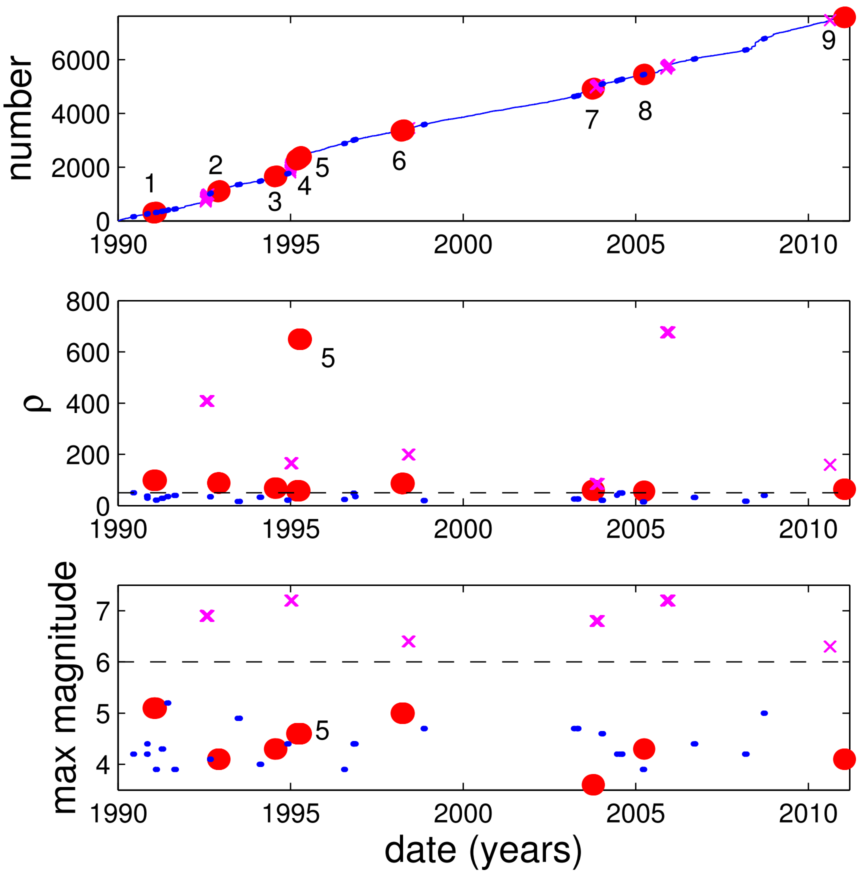
Figure 7. Temporal distribution of the 19 loading transients found with $\ell = 80$ km and $\tau = 100$ days. Legend is the same as in Figure 3. We here used a threshold $\rho > 35$ to display the most intense transients (in red). The 8 most intense transients are labelled from 1 to 8 in the upper left graph, with the two occurring in the north having yellow labels.

Figure 8. Map of the large-scale transients. The red dots show the earthquakes that are part of the 8 most intense transients shown in Figure 7.

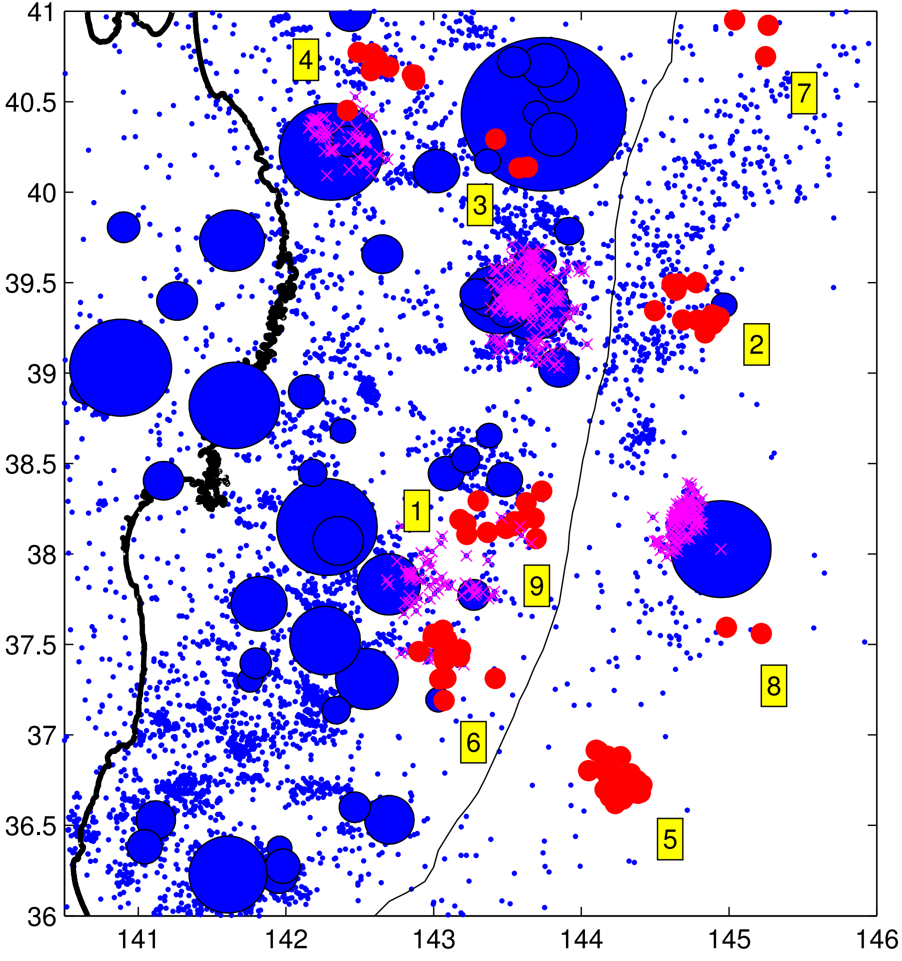
Figure 9. Occurrence times of the four intense transients in the southern part of the zone, after grouping the two tightly clustered pairs 4-5 and 6-7 together to form two single transients. The best linear fit for the starting dates of these four transients is shown with the black line, and gives a recurrence time of 5.9 years.

Figure 10. Year 1998 subduction transient, characterized by the maximum ρ of all transients in Figure 8. The area covered by the transient is shown in yellow. Top right graph: cumulative activity in the area of the transient, from 1996 to 2002, with the best linear fit, in black, for these 6 years. The duration of the transient is shown in red. Inset: background rate (in number of events per year) estimated for the area of the transient. Bottom right graph: magnitude vs time for the earthquakes in the transient, from 1996 to 2002.

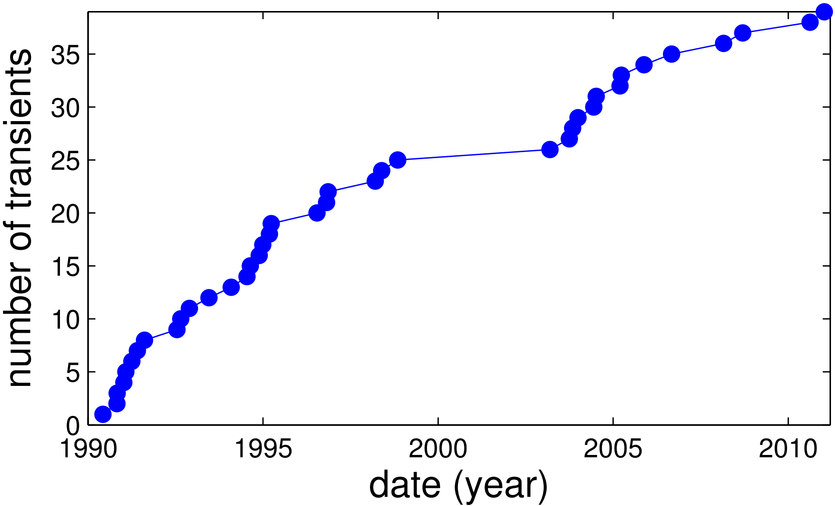

41
40.5
40
39.5
39
38.5
38
37.5
37
36.5
36
141
142
143
144
145
146
11
10
9
8
7
6
5
4
3
2
1
number of m3.5+ earthquake/year/100x100 km^2



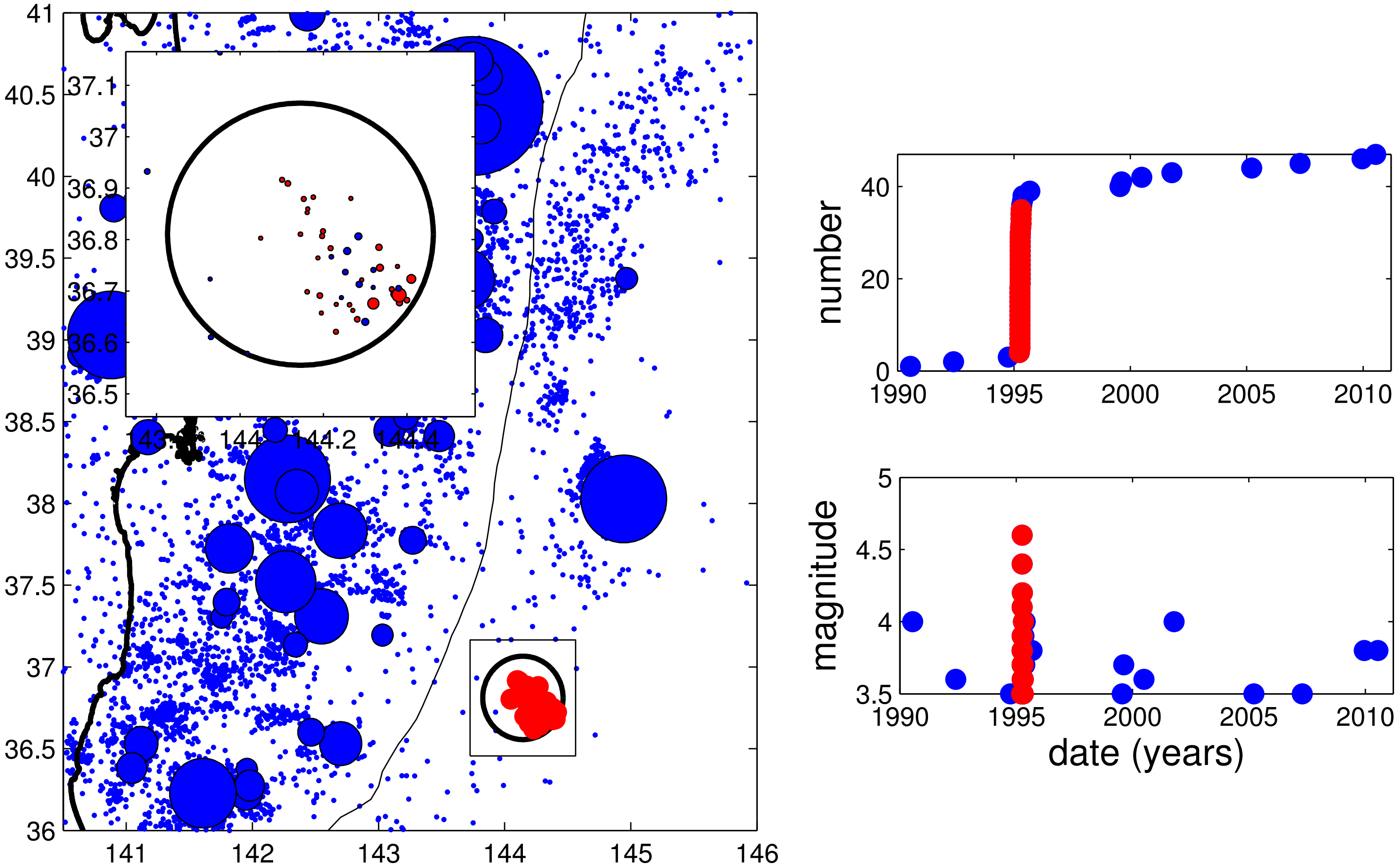
number
6000
4000
2000
0
1990
1995
2000
2005
2010
1
2
3
4
5
6
7
8
9
800
600
400
200
0
ρ
5
max magnitude
7
6
5
4
date (years)



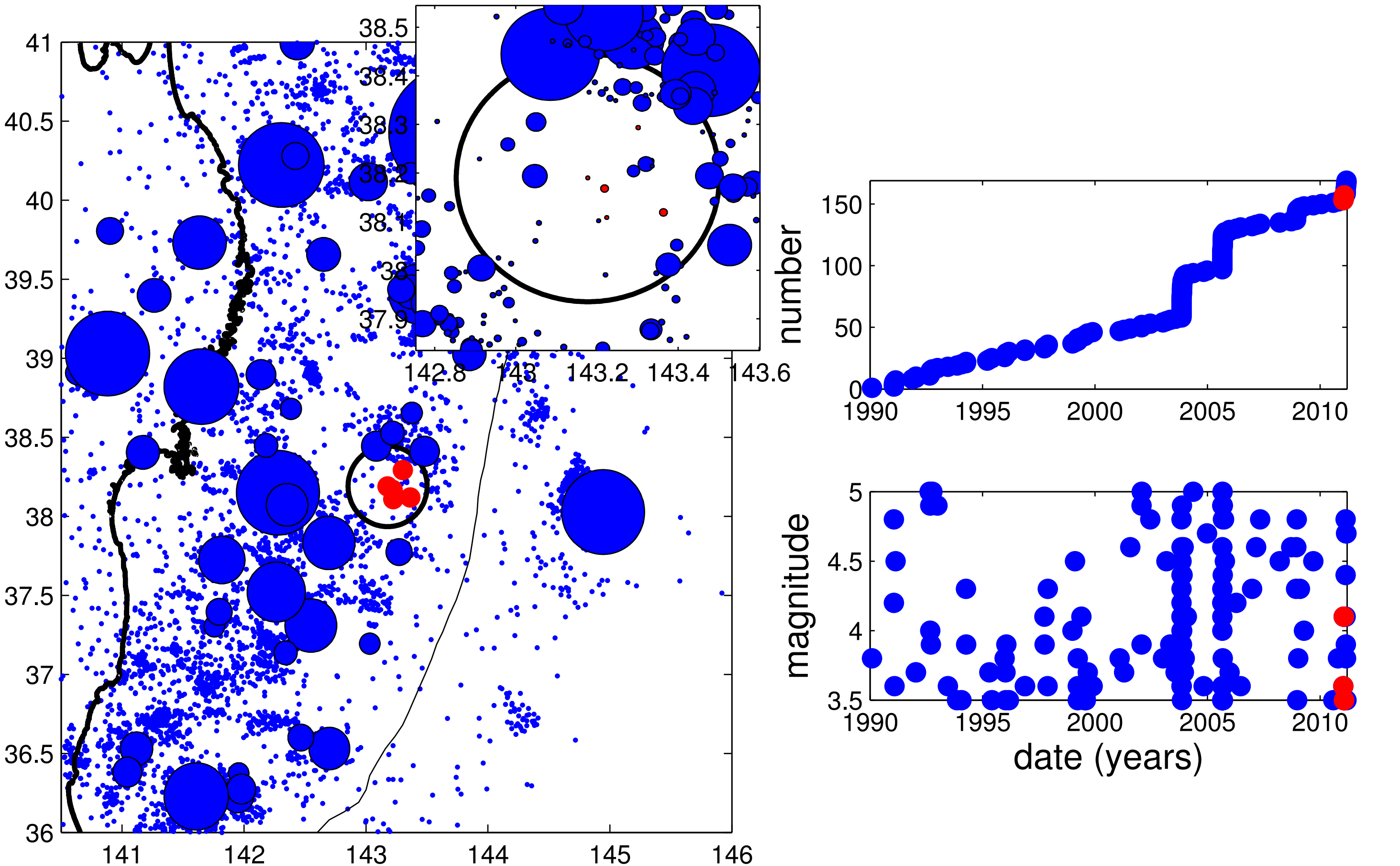
41
40.5
40
39.5
39
38.5
38
37.5
37
36.5
36
141
142
143
144
145
146
1
2
3
4
5
6
7
8
9



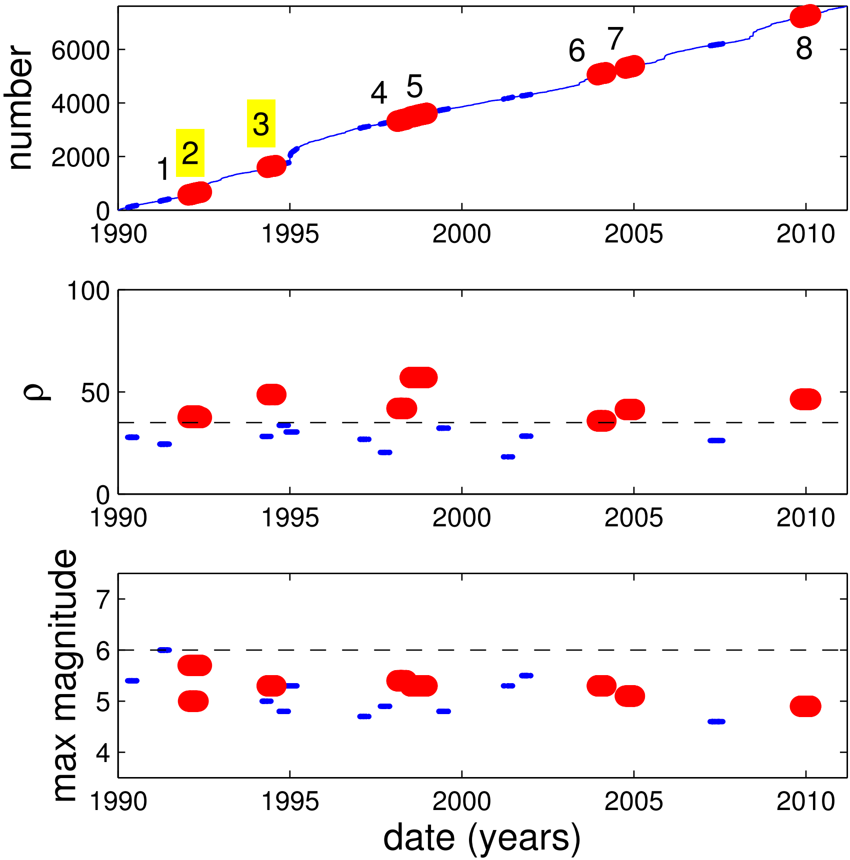
number of transients
0
5
10
15
20
25
30
35
1990
1995
2000
2005
2010
date (year)



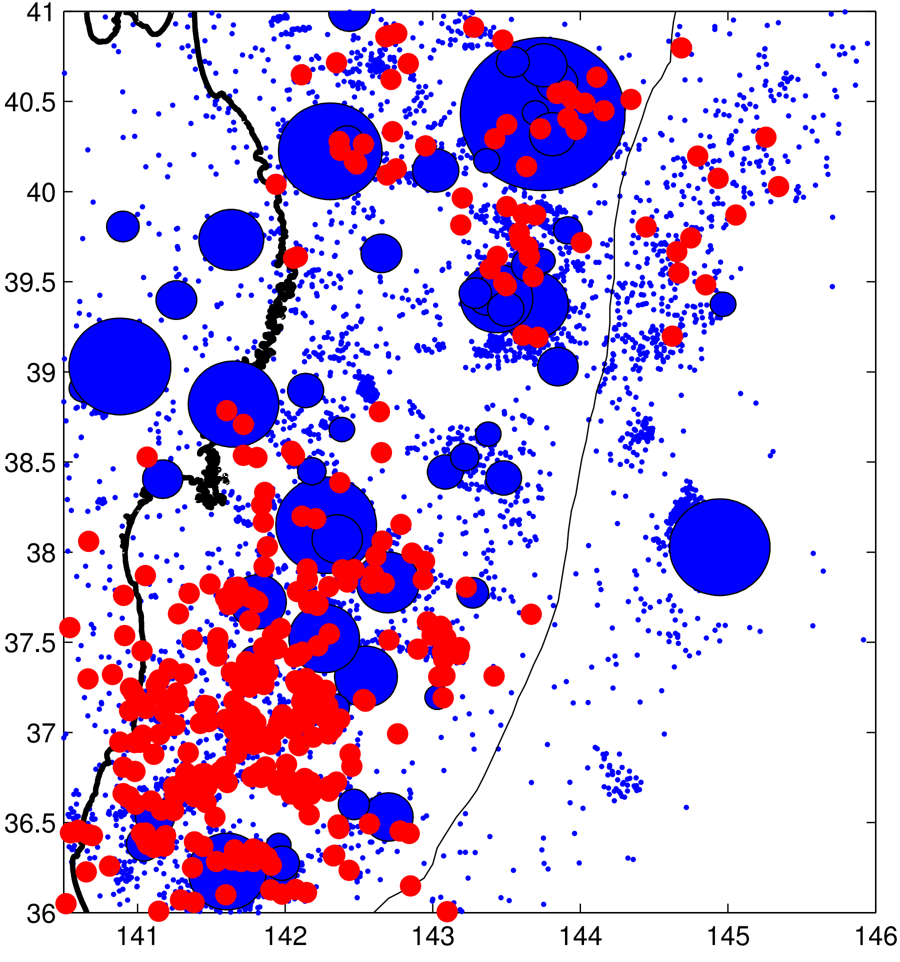
41
40.5
40
39.5
39
38.5
38
37.5
37
36.5
36
141
142
143
144
145
146
37.1
37
36.9
36.8
36.7
36.6
36.5
143.8
144
144.2
144.4
number
0
20
40
1990
1995
2000
2005
2010
magnitude
3.5
4
4.5
5
1990
1995
2000
2005
2010
date (years)

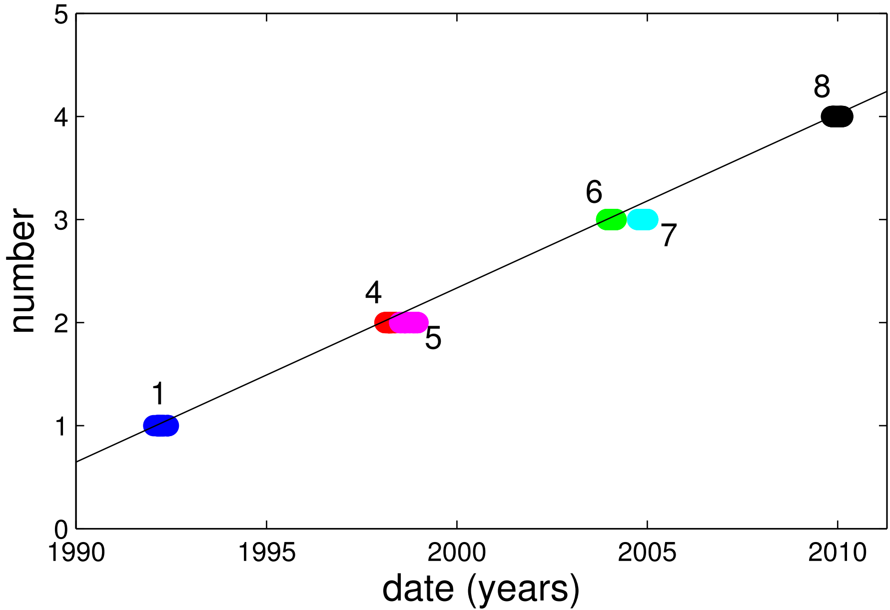


number
magnitude
date (years)

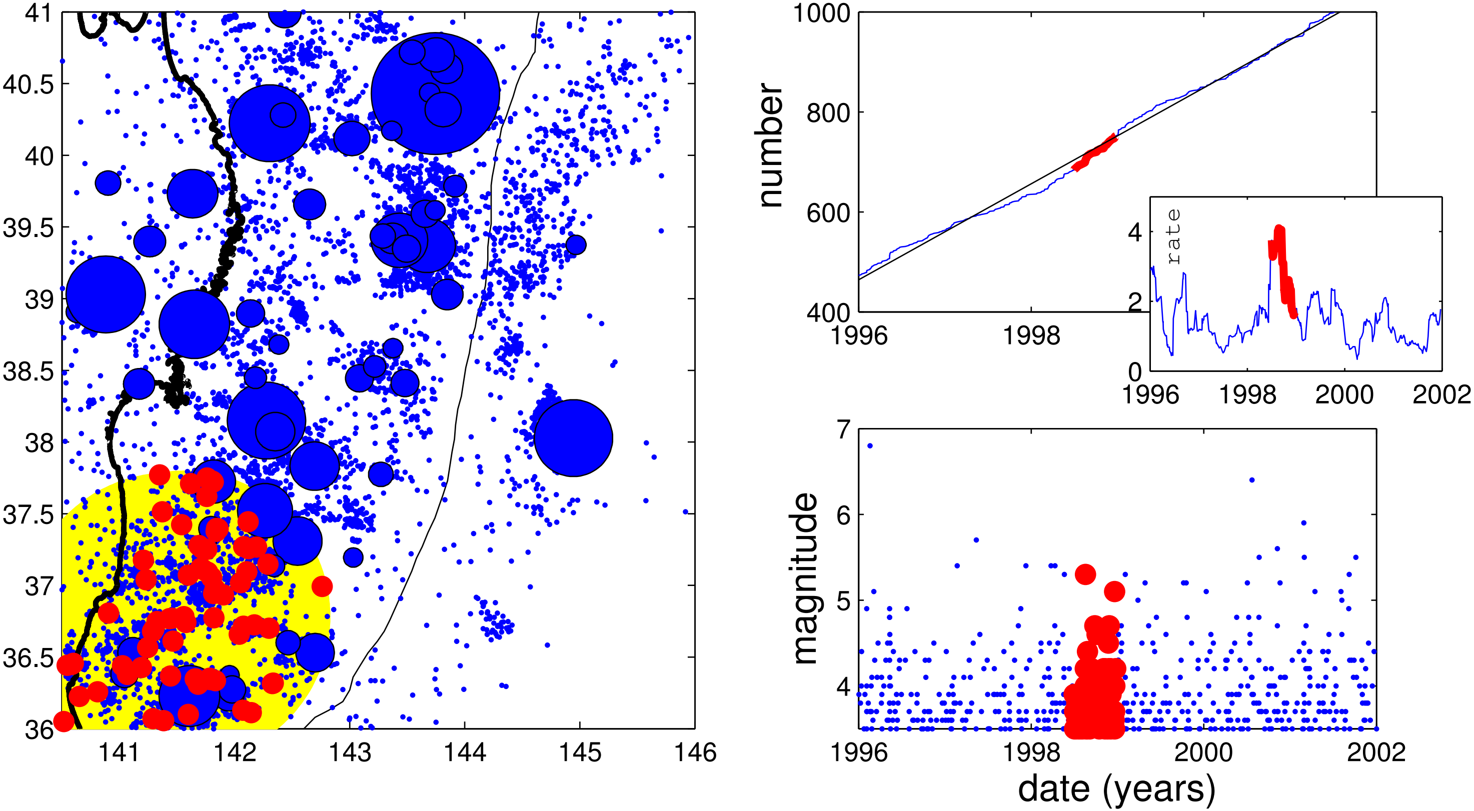


number
6000
4000
2000
0
1
2
3
4
5
6
7
8
1990
1995
2000
2005
2010
ρ
100
50
0
max magnitude
7
6
5
4
date (years)


41
40.5
40
39.5
39
38.5
38
37.5
37
36.5
36
141
142
143
144
145
146



number
date (years)
0
1
2
3
4
5
1990
1995
2000
2005
2010
1
4
5
6
7
8



41
40.5
40
39.5
39
38.5
38
37.5
37
36.5
36
141
142
143
144
145
146
number
1000
800
600
400
1996
1998
rate
4
2
0
2000
2002
magnitude
7
6
5
4
date (years)