Detection and characterization of transient forcing episodes affecting earthquake activity in the Aleutian Arc system

T. Reverso a,c,* , D. Marsan a,c, A. Helmstetter b,c

a ISTerre, Univ. de Savoie, F-73376 Le Bourget du Lac, France
b ISTerre, Univ. Grenoble Alpes, F-38041 Grenoble, France
c CNRS, ISTerre, F-73376 Le Bourget du Lac, France

A R T I C L E   I N F O

Article history:
Received 5 August 2014
Received in revised form 2 December 2014
Accepted 4 December 2014
Available online 31 December 2014
Editor: P. Shearer

Keywords:
seismic swarms
crustal deformation
creep

A B S T R A C T

Crustal, slow deformation transients can be caused by fluid or magmatic intrusions, and by slow slip on faults. They can affect earthquake dynamics, if they occur close to or within seismically active zones. We here further develop, and test, a statistical method for detecting and characterizing seismicity anomalies that is only based on earthquake occurrence times and locations. We make use of this method to analyze the 2004–2013 seismicity at m c = 3.5 in the Aleutian subduction system, to find six statistically significant anomalies, with typical 1 day duration and 30 to 50 km size, that are likely related to slow deformation transients. They tend to be located in zones characterized by intermediate seismic coupling, and to mark the termination of past large to mega-thrust earthquakes. These anomalies account for a non-negligible (9%) part of the total activity, proving that non-stationary aseismic loading plays an important role in the dynamics of crustal deformation.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Earthquakes occur as a consequence of accumulating stress in the crust. Estimating the rate at which stress loads a fault is a particularly challenging task, as systematic in situ measurement at seismogenic depth is still out of reach. Monitoring seismicity rates λ(x, y, t), i.e., the number of earthquakes per unit time and unit area/volume at location (x, y) and time t, as proxies of stressing rates is a common approach, but it implies modeling how these two quantities relate to one another. Mechanical modeling of earthquake nucleation, e.g., using dislocation and friction models, generally accounts for loading due to long-term tectonic stressing, plus stress changes imparted by seismic sources big enough so that their characteristics are known with good certainty (Stein, 1999). However, it has been evidenced that small, poorly characterized sources also contribute significantly to the dynamics of seismicity (Helmstetter et al., 2005; Marsan, 2005; Meier et al., 2014). Stochastic modeling thus offers an alternative approach, that fully uses the seismicity information at hand, albeit at the cost of simplifying assumptions, in particular that earthquakes of equal magnitude behave the same as triggers. The seismicity rate λ here results from the two distinct contributions of a background aseismic rate μ and a seismic rate ν that can be modeled from the past history of earthquake occurrences:

λ(x, y, t) = μ(x, y, t) + ν(x, y, t).

Recent developments in seismology have emphasized the ubiquity of stress loading contributions from aseismic (i.e., not involving rupturing at seismic velocities), local processes, including silent fault slip within and underneath the seismogenic layer (Schwartz and Rokosky, 2007; Peng and Gongberg, 2010). Episodes of aseismic loading can thus cause changes in seismicity dynamics, if they occur in the proximity of faults close enough to failure. Transient deformation, or therein after ‘transients’, can therefore be revealed by the occurrence of seismic swarms, which do not obey usual mainshock–aftershock patterns.

Studies aimed at detecting transients with stochastic methods have focused on specific sequences, typically at the scale of tens of kilometers (Hainzl and Ogata, 2005; Llenos et al., 2009; Llenos and McGuire, 2011; Daniel et al., 2011; Peng et al., 2012). They considered that the relative evolution of the loading rate μ is the same at all points of the system, and thus decoupled μ(x, y, t) = μs(x, y) × μt(t) to invert for the marginal μs. A methodological framework for performing this inversion is described in Marsan et al. (2013a). At the regional scale, from 100 to 1000 km, swarms only cover a small portion of the seismically active surface/volume. It is then inappropriate to consider that μ follows the same evolution everywhere. Transient aseismic loading
must then be modeled as local, both in space and time, and the decoupling proposed in previous studies must be relaxed.

Preliminary attempts at doing so by Marsan et al. (2013b) were motivated by the question as to whether the swarm activity preceding the 2011 M9.0 Tohoku earthquake was unique or not to this part of the Japanese subduction. The method then developed however provides only a partial account of the significance of the estimated changes in background rate, through the computation of the Akaike Information Criterion (Akaike, 1973). Other approaches have been based on visual inspection of seismicity patterns (Holtkamp and Brudzinski, 2011), or on clustering criteria probing swarm occurrences at specific spatial and temporal scales (Vidale and Shearer, 2006). Zaliapin and Ben-Zion (2013a, 2013b) developed the nearest-neighbor method of Zaliapin et al. (2008) to discriminate swarm activity, using a priori fixed model parameters.

We extend the approach of Marsan et al. (2013b) by fully measuring the significance level of suspected episodes of aseismic deformation: tests of the method are then run to evaluate its accuracy and resolution power. The analysis of regional seismicity in the Aleutian arc is then performed, to compare the inverted transients with independent accounts of aseismic transients, and to investigate the spatial distribution of these deformation episodes.

2. Method

We define the seismicity as the combination of two components. The first is the seismicity due to aseismic processes, including tectonic loading. This spontaneous seismicity is not triggered by precursory events and is called the background seismicity. The second term corresponds to aftershocks, i.e. earthquakes triggered by previous shocks. Hereinafter, we assume that this triggering can be modeled by empirical laws (i.e. productivity law, Omori’s law (Utsu, 1961; Omori, 1894)). In our approach, the seismicity associated with episodic aseismic phenomena, like deformation transients, can be modeled as an increase in the rate of background activity since it is not triggered by previous earthquakes. The aim is thus to evaluate the spatio-temporal variations of the background seismicity \( \mu(x, y, t) \), which embodies both constant tectonic loading and loading through episodic aseismic processes (e.g., fluid intrusions or slow slip events). The latter cause \( \mu \) to fluctuate in time, unlike the tectonic loading which is assumed to be constant in rate at the time scale of instrumental earthquake catalogs.

The overall approach follows and further develops the method of Marsan et al. (2013b). Two models are optimized against the data, (1) the null-hypothesis model \( M_0 \), in which the background activity is only caused by tectonic loading, hence a constant but spatially variable \( \mu(x, y) \), and (2) model \( M_1 \) in which the background activity also includes time-fluctuating processes, hence allowing \( \mu(x, y, t) \) to also vary in time. The two models are then compared using a Monte-Carlo method, to search for significant episodes of changes in background rate, hence of slow, aseismic deformation. Model \( M_0 \) is the null hypothesis of no changes in background rate. We now detail the method, which can be divided into 3 steps.

2.1. Null hypothesis: model \( M_0 \)

We use the space–time ETAS model, which represents earthquakes as points occurring with rate-density \( \lambda_0(x, y, t) \), defined as the mean number of earthquakes per unit area and unit time. This rate is the sum of two terms:

\[
\lambda_0(x, y, t) = \mu(x, y) + \nu(x, y, t)
\]

(1)

with \( \mu(x, y) \), the background seismicity, assumed to be constant in time in this first step, and \( \nu(x, y, t) \) a term of interactions between earthquakes. The latter term is defined as the product of a temporal and a spatial influence

\[
\nu(x, y, t) = \frac{\kappa(m)}{(t + c)^{\mu}} \frac{(\gamma - 1)L(m)^{\gamma - 1}}{2\pi(x^2 + y^2 + L(m)^2)^{\gamma + 1/2}}
\]

where \( c, \gamma \) and \( p \) are constants, \( L(m) \) and \( \kappa(m) \) represent the rupture length and the productivity law, respectively (Ogata, 1988; Zhuang and Chang, 2005). The productivity law \( \kappa(m) \) is defined as

\[
\kappa(m) = \kappa_0 \times e^{\alpha(m - m_0)}
\]

where \( \kappa_0 \) and \( \alpha \) are constant and \( m_0 \) is the magnitude threshold. We assume that the rupture length \( L \) scales with magnitude according to

\[
L(m) = L_0 10^{0.5(m - m_0)}
\]

where \( L_0 \) is the rupture length for an earthquake of magnitude \( m_0 \).

2.1.1. Smoothing

This first model thus requires 6 parameters \( \theta = \{\alpha, p, c, L_0, \gamma, \alpha \kappa_0\} \). The probability \( \omega_i \) that earthquake \( i \) is a background earthquake is

\[
\omega_i = \frac{\mu_i}{\mu_i + v_i}
\]

(3)

where \( \mu_i \) and \( v_i \) are respectively the background seismicity and the interaction term for earthquake \( i \).

We estimate the background intensity \( \mu(x, y) \) by smoothing these probabilities (Zhuang et al., 2002) over all earthquakes \( i \):

\[
\mu(x, y) = \frac{1}{T} \sum \omega_i Z_L(x - x_i, y - y_i)
\]

(4)

where \( T \) is the duration of the catalog and \( Z_L(x - x_i, y - y_i) \) is defined as

\[
Z_L(x - x_i, y - y_i) = \frac{1}{2\pi L^2} e^{-\frac{(x - x_i)^2 + (y - y_i)^2}{L^2}}
\]

(5)

with \( L \) a smoothing length.

2.1.2. ETAS parameter estimations and inversions

To optimize the model, we maximize the log-likelihood defined as

\[
\ell(\theta) = \sum \left[ \ln \lambda_0(x_i, y_i, t_i) - \int_0^T \int_0^L \lambda_0(x, y, t) dx dy dt \right]
\]

(6)

where the integral is performed over the total duration \( T \) and surface \( S \) of the area containing the target earthquakes.

To optimize model \( M_0 \), we follow the method of Zhuang et al. (2002). We start with a uniform background rate \( \mu(x, y) \) with an arbitrary positive value. Given this \( \mu(x, y) \), the best parameters are searched by maximizing \( \ell(\theta) \). The background probabilities \( \omega_i \) are then computed, and smoothed according to Eq. (5) to yield an updated \( \mu(x, y) \). Then the best parameters \( \theta \) given this new \( \mu(x, y) \) are computed, and so on until convergence of the solution, both for \( \theta \) and \( \mu \). This solution does not depend on the initial choice of \( \mu(x, y) \), but does depend on the smoothing length \( L \).

To track possible temporal changes in the background rate, we use a discretized version of \( \mu \). We define a regular grid in time and space, each cell having a space area \( L \times L \) and a duration \( T \). The stationary background rate of cell \( i \) with center \( [x_i, y_i] \) is therefore...
\[\mu_0^{(i)} = \frac{1}{L^2} \int_{[x-X_i]<L/2} \int_{[y-Y_i]<L/2} \mu(x, y) \, dx \, dy.\]  
(7)

We define the log-likelihood for model \(M_0\) and for cell \(i\) as
\[\ell_0^{(i)} = -\mu_0^{(i)} \tau L^2 + N^{(i)} + \sum_j \ln(\mu_0^{(i)} + v_j)\]  
(8)

where
\[N^{(i)} = \int_{[x-X_i]<L/2} \int_{[y-Y_i]<L/2} v(x, y) \, dx \, dy \, dt\]  
(9)

and the summation is on all earthquakes \(j\) occurring in cell \(i\) centered on \([X_i, Y_i]\) and on time \(T_i\).

### 2.2. Time-varying background rate: model \(M_1\)

In model \(M_1\), we relax the hypothesis of a stationary background rate. Each cell \(i\) is now characterized by a background rate-density \(\mu_1^{(i)}\) and the log-likelihood \(\ell_1^{(i)}\) is given by
\[\ell_1^{(i)} = -\mu_1^{(i)} \tau L^2 - N^{(i)} + \sum_j \ln(\mu_1^{(i)} + v_j)\]  
(10)

In this computation, the terms \(N^{(i)}\) and \(v_j\) are the same as for \(\ell_0\) defined by (8); the influence of other, past earthquakes on the earthquakes contained in cell \(i\) is the same for both models.

### 2.3. Models comparison and significance

To evaluate the difference between models \(M_1\) and \(M_0\), we calculate the log-likelihood gain \(\Delta \ell^{(i)}\) for each cell \(i\):
\[\Delta \ell^{(i)} = \ell_1^{(i)} - \ell_0^{(i)}\]  
(11)

We finally test whether the obtained values of \(\Delta \ell\) can be found by chance, if model \(M_0\) is true. This amounts to check whether model \(M_1\) is significantly better than the null hypothesis \(M_0\). To do this, we use a Monte–Carlo method and construct 1000 synthetic catalogs generated with ETAS parameters and \(\mu(x, y)\) estimated from the data, see Section 2.1. These synthetic catalogs have a stationary background, and therefore do not contain any transient episode. We then compute \(\Delta \ell\) for all cells and for each synthetic catalog. We finally compare the values of \(\Delta \ell\) obtained for the real data and for the synthetic ETAS catalogs. This allows us to quantify the probability \(p(\Delta \ell)\) that a temporal change in background rate is significant, i.e., is anomalous compared with the values of \(\Delta \ell\) obtained for the synthetic catalogs. More precisely, \(p(\Delta \ell)\) is defined as:
\[p(\Delta \ell) = \frac{1}{1000} \sum_{n=1}^{1000} H(\max \Delta \ell^{(n)} < \Delta \ell)\]  
(12)

where \(H(x) = 1\) if \(x\) is true, 0 otherwise, and \(\max \Delta \ell^{(n)}\) is the maximum value of \(\Delta \ell\) for the synthetic catalog number \(n\). We have run tests in order to assess the capacity and the shortcomings of the method, see Appendix A. These tests show more particularly that the detection parameters \(L\) and \(\tau\) are important, as only transients of similar sizes and durations can be detected. Moreover, we find that anomalous aftershock sequences can sometimes appear as transients. This can be caused by the fact that some mainshocks produce stronger aftershock sequences than expected by the model. As we are mostly interested by swarms unrelated to large shocks, we discuss at the end of next section how these anomalous aftershock sequences detected as transients can be separated from the other transients.

### 3. Deformation transients in the Aleutian arc system

Subduction zones exhibit diverse sliding modes (seismic or aseismic, e.g. Schwartz and Rokosky, 2007), mainly owing to a heterogeneous seismic coupling. Low coupling areas are potential candidates for slow deformation episodes, as for example observed in the Boso, Japan (Ozawa et al., 2007). The Aleutian arc is characterized by clear spatial variations in coupling (Scholz and Campos, 2012), as well as hosting both mega-thrust earthquakes up to \(M_w\) 9.2 and slow slips events (Ohita et al., 2006) in instrumental times. Holtkamp and Brudzinski (2011) identified 5 transients after 2004 in the arc, listed in Table S2 in the supplementary material. These transients lasted for almost 2–3 days and they have spatial extents between 400 and 2000 km².

#### 3.1. Data

We use the PDE seismicity catalog, available at the USGS data center, for the Aleutian arc from 01/01/2004 to 31/12/2013 (http://www.earthquake.usgs.gov/earthquakes/search/), for latitudes from 48° to 60°N, longitudes from 165° to −145°E and depth smaller than 100 km. We estimate the magnitude of completeness at \(m_c = 3.5\) for this set. We thus end up with \(N = 3598\) earthquakes. Fig. 1 represents the seismicity in the study area. We do not further use the depths of the earthquakes in the following treatment, and only compute epicentral distances, because of the large uncertainties in depth estimates.

#### 3.2. Deformation transients

Our method identifies transients at specific scales chosen by the operator. We first investigate the scales of the four transients identified by Holtkamp and Brudzinski (2011) (Table S2 in supplementary material), and thus fix \(L = 30\) km and \(\tau = 1\) day. The ETAS parameters estimated for \(L = 30\) km are: \(\alpha = 1.079, p = 1.021, c = 0.0044\) day, \(\gamma = 2.758, L_0 = 1.701\) km for \(m_0 = 3.5, \kappa_0 = 0.015\). We represent in Fig. S3 in supplementary material the obtained anomalies and we sum up the results in Table 1.

We identify eight anomalies with probability to be anomalous >70%, among which four are effectively transients: they correspond to an increase of background seismicity not directly related to an aftershock sequence, see Table 1. These 4 transients are located in different areas and take place at different times but their seismic moments are roughly similar, between 1 and \(8 \times 10^{17}\) N m, corresponding to moment magnitudes between 5.3 and 5.9. Only one of them (A3) is listed in Holtkamp and Brudzinski (2011).

As discussed in Section 3.3, our method relies on ensemble averaged laws describing the number of aftershocks triggered by mainshocks. Natural variability around these laws can generate large fluctuations in these numbers. Since this variability is not accounted for by the model, it ends up fitting very productive aftershocks sequence by an ad-hoc increase in background rate, potentially causing the detection of an anomaly. We examined each anomaly, searching for possible large shocks preceding them that could explain their occurrence. This visual inspection led to the rejection of 4 out of the 8 anomalies, see Table 1. Changing the detection parameters \(L\) and \(\tau\) leads to the detection of other anomalies, see Table 2. We detect these anomalies for other scales but with different probabilities (cf. Table S4). Among the four transients listed in Holtkamp and Brudzinski (2011) (i.e., after merging their last two transients which effectively overlap), we only find three. The last event (i.e., A10), located on Kodiak Island, is undetected because its probability to be a transient is less than 70%. We detect one extra other transients that were not identified by Holtkamp and Brudzinski (2011), and two others after 2010.
Fig. 1. Seismicity in the Aleutian arc between 2004 and 2013. (a) Epicenter locations. The trench (Hayes et al., 2012) is shown with the dashed black line, and volcanoes with white triangles. (b) Cumulative number and magnitude of earthquakes as a function of time.

Table 1
Anomalies with a probability to be significant >70% for $L = 30$ km and $\tau = 1$ day. The separation into actual transients and aftershock sequence is done visually, see text for explanations.

<table>
<thead>
<tr>
<th>Name</th>
<th>Starting date (yr)</th>
<th>Lat. start</th>
<th>Lon. start</th>
<th>Probability to be anomalous</th>
<th>Transient like?</th>
<th>Number of Eqs.</th>
<th>Seismic moment (Nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_1$</td>
<td>2005.450</td>
<td>50.946</td>
<td>179.485</td>
<td>82%</td>
<td>Transient</td>
<td>19</td>
<td>$5.66 \times 10^{17}$</td>
</tr>
<tr>
<td>$A_2$</td>
<td>2006.354</td>
<td>52.296</td>
<td>-169.469</td>
<td>96%</td>
<td>Transient</td>
<td>8</td>
<td>$4.43 \times 10^{18}$</td>
</tr>
<tr>
<td>$A_3$</td>
<td>2006.498</td>
<td>50.946</td>
<td>-179.657</td>
<td>97%</td>
<td>Transient</td>
<td>14</td>
<td>$8.20 \times 10^{17}$</td>
</tr>
<tr>
<td>$A_4$</td>
<td>2008.598</td>
<td>52.060</td>
<td>-175.690</td>
<td>99.99%</td>
<td>Transient</td>
<td>31</td>
<td>$6.21 \times 10^{17}$</td>
</tr>
<tr>
<td>$A_5$</td>
<td>2009.781</td>
<td>52.565</td>
<td>-167.178</td>
<td>82%</td>
<td>Aftershock sequence</td>
<td>12</td>
<td>$1.08 \times 10^{19}$</td>
</tr>
<tr>
<td>$A_6$</td>
<td>2010.543</td>
<td>52.565</td>
<td>-169.849</td>
<td>99%</td>
<td>Aftershock sequence</td>
<td>37</td>
<td>$1.48 \times 10^{18}$</td>
</tr>
<tr>
<td>$A_7$</td>
<td>2010.841</td>
<td>50.946</td>
<td>-176.610</td>
<td>99.99%</td>
<td>Transient</td>
<td>17</td>
<td>$1.36 \times 10^{17}$</td>
</tr>
<tr>
<td>$A_8$</td>
<td>2011.670</td>
<td>51.756</td>
<td>-171.781</td>
<td>99%</td>
<td>Aftershock sequence</td>
<td>19</td>
<td>$1.64 \times 10^{17}$</td>
</tr>
</tbody>
</table>

Table 2
Table showing anomalies with a probability to be a transient >70% for $L = 50$ km and $\tau = 1$ day.

<table>
<thead>
<tr>
<th>Name</th>
<th>Starting date (yr)</th>
<th>Lat. start</th>
<th>Lon. start</th>
<th>Probability to be anomalous</th>
<th>Transient like?</th>
<th>Number of Eqs.</th>
<th>Seismic moment (Nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_1$</td>
<td>2005.450</td>
<td>50.767</td>
<td>179.486</td>
<td>96%</td>
<td>Transient</td>
<td>30</td>
<td>$5.84 \times 10^{17}$</td>
</tr>
<tr>
<td>$A_2$</td>
<td>2008.8319</td>
<td>51.216</td>
<td>-177.749</td>
<td>80%</td>
<td>Transient</td>
<td>15</td>
<td>$3.17 \times 10^{17}$</td>
</tr>
<tr>
<td>$A_{10}$</td>
<td>2009.2415</td>
<td>56.163</td>
<td>-152.967</td>
<td>72%</td>
<td>Transient</td>
<td>7</td>
<td>$1.24 \times 10^{18}$</td>
</tr>
<tr>
<td>$A_6$</td>
<td>2010.543</td>
<td>52.565</td>
<td>-170.126</td>
<td>99.8%</td>
<td>Aftershock sequence</td>
<td>22</td>
<td>$1.25 \times 10^{19}$</td>
</tr>
<tr>
<td>$A_7$</td>
<td>2010.841</td>
<td>50.767</td>
<td>-176.942</td>
<td>96%</td>
<td>Transient</td>
<td>19</td>
<td>$1.69 \times 10^{17}$</td>
</tr>
</tbody>
</table>

The inverted $\alpha$ value of 1.079 (at $L = 30$ km), while typical of regional ETAS studies (cf. Marsan et al., 2013a on this topic), is low compared to $\alpha$ values obtained with windowing methods (e.g., Helmstetter et al., 2005). We however checked that the six transients found with $\alpha = 1.079$ were effectively also found when imposing $\alpha = 2.3$ all throughout the treatment.

3.3. Comparison with the Nearest-Neighbor earthquake Distance method (NND)

Zaliapin and Ben-Zion (2013a, 2013b) developed a method, called the “Nearest-Neighbor earthquake Distance method” (NND), in order to group earthquakes in clusters and also to distinguish background activity from earthquake clusters. Their method is
based on estimating the so-called “nearest” earthquake for each event. The nearest earthquake, or parent of event $j$, is defined as the earthquake $i$ with the smallest spatio-temporal distance $\eta_{ij}$

$$\min_{i}(\eta_{ij}) = \min_{i}(T_{ij} \times R_{ij})$$

where $T_{ij}$ is the normalized temporal distance and $R_{ij}$, the normalized spatial distance between earthquakes

$$T_{ij} = t_{ij} \times 10^{-q \eta_{ij}}, \quad R_{ij} = (r_{ij})^{d_f} \times 10^{-q \eta_{ij}}$$

with $t_{ij}$ the inter-occurrence time between the two earthquakes, $r_{ij}$ their epicentral distance, $d_f$ the fractal dimension of the set of earthquake epicenters, $b$ the $b$-value of the Gutenberg–Richter law and $q$ a constant fixed to $q = 0.5$ that account for the scaling of rupture length with magnitude. For each earthquake $j$, the nearest earthquake $i^*$ is the event with the smallest spatio-temporal distance $\eta_{i^*,j}$, corresponding to a distance $R_{i^*,j} = R_{i,j}$ and a time interval $T_{i^*,j} = T_{i,j}$. Background earthquakes are those with large $\min_{i}(\eta_{ij})$, and can be extracted from the rest by separating the distribution $R_{i,j}^t$ vs. $T_{i,j}^t$ into a “clustered” and a “background” part.

We apply their method to the Aleutian arc using $b = 1$, $q = 0.5$ and $d_f = 1.6$. Results are shown in Fig. 2. The distance to the nearest earthquake is characterized by a bi-modal distribution separated by the limit $\log R + \log T = -5$ (i.e., $\eta > \eta_0 = 10^{-5}$). The earthquakes with $\eta > \eta_0$ are supposed to represent the Poissonian part of the background seismicity. It is important to emphasize that the definition of background seismicity by Zaliapin and Ben-Zion (2013a) differs from ours: in their approach, background activity can only result from a stationary Poisson process. In our case, all activity not triggered by previous earthquakes is background, thus including possible non-stationary transients. As such, the episodes of transient, slow deformation we are seeking are part of the “clustered” group as defined by Zaliapin and Ben-Zion (2013a).

In order to extract background transients (our sense), we follow Zaliapin and Ben-Zion (2013a). Binary linking all the earthquakes to their nearest-neighbor defines distinct clusters. For each cluster, we compute the number of leaves $N_L$ and the average leaf depth ($d$) (Zaliapin and Ben-Zion, 2013a). Plotting ($d$) as a function of $N_L$, as illustrated in Fig. S5 in supplementary material, should allow to single out transients. Applying this procedure to our dataset, we observe that transients identified with our method have large ($d$) values, and low $N_L$ (see Fig. 3).

We search in the NND clusters those that contain our transients as listed in Table 2. We observe that our transients are effectively located in the “swarm-like” part of the graph, while $A_L$, identified previously as being part of an aftershock sequence, is here again clearly in the “aftershock” domain. Using the NND method therefore can be used in order to suppress the aftershock sequences from the anomalies detected by our method. Conversely, using the NND method alone gives “swarm-like” clusters that are not anomalous according to our triggering model (see Fig. 3). Moreover, the size and duration of the NND clusters containing our transients are generally very different from the size and duration of the corresponding transients (Table S6 in supplementary material). For example, the transient $A_1$ is detected with $L = 30$ km and $\tau = 1$ day, and with $L = 50$ km and $\tau = 1$ day, while the corresponding NND cluster has a duration of almost 3 years and a size of $180 \times 180$ km$^2$.

4. Discussion

We found 6 transients ($A_1$, $A_3$, $A_4$, $A_7$, $A_9$, $A_{10}$) that are not related to aftershock sequences between 2004 and 2013 (see Fig. 5). Three of them ($A_3$, $A_9$, $A_{10}$) are also detected by Holtkamp and Brudzinski (2011). We moreover found two others in the same period ($A_1$ and $A_4$), and another one after 2009 ($A_7$). These transients are not located close to volcanoes, except $A_4$, which began in August 2008 and is located underneath Kasatochi strato-volcano (52.169°N, −175.511°E) at a depth of 5 to 10 km, see Fig. 5. This transient corresponds to a swarm that occurred along with the eruption of Kasatochi volcano on August 7, 2008 (Ruppert et al., 2011), and lasted for about 3 days.

To investigate whether the five other transients are on the subduction interface, we use the Alaska Earthquake Information Center (AEIC) catalog to better constrain the depth (see Figs. S7 to S12 in supplementary material). The focal mechanisms taken from the Harvard CMT catalog effectively suggest that the transients $A_1$, $A_3$, $A_7$, $A_9$, and $A_{10}$, are located on the subduction interface, see Fig. 4.
We then estimate the depths by interpolating the slab iso-depth contours of Hayes et al. (2012), which range from 10 km (A10) to 30 km (A1, A3) and possibly even 40 km (A9). These depths extend past the 20 km transition depth proposed by Delahaye et al. (2009) to separate aseismic transients that do (up-dip) and do not (down-dip) trigger seismicity.

We did not find any transient in the outer rise, unlike in Japan where swarms probably triggered by upwelling fluid intrusions in the extensive domain of outer rise were found by Marsan et al. (2013a).

Slow Slip Events (SSE) can be found in subduction zones and are thought to be a characteristic of low seismic coupling zones, although, the very obvious Tohoku foreshock swarm that lasted for about 1.5 months in January–February 2011 is very likely related to slow slip in an otherwise strongly coupled zone (Kato et al., 2012; Marsan et al., 2013a).

The Aleutian Arc has laterally-heterogeneous coupling, with locked zones (e.g., Kodiak Island) as well as creeping zones characterized by aseismic phenomena (Cross and Freymueller, 2007; Freymueller et al., 2013; Scholz and Campos, 2012). Non-Volcanic Tremor (NVT) and Deep Low-Frequency Earthquakes (DLFE) are found along the arc (Peterson et al., 2011; Gomberg and Prejean, 2013) while several instances of SSE are localized in southern Alaska (Ohta et al., 2006; Peterson and Christensen, 2009; Fu and Freymueller, 2013).

Transient A10 is located beside Kodiak Island, relatively close to the surface: a mean depth of 10 km is suggested by the slab iso-depth contours (see Fig. S12 in supplementary material). A broad zone of 200 km × 150 km centered on Kodiak Island is expected to be locked (Freymueller et al., 2013). This zone acted as the second asperity that failed during the Mw 9.2 1964 great Alaskan earthquake, with a local co-seismic slip of about 20 m (Suito and Freymueller, 2009). This transient, also detected by Holtkamp and Brudzinski (2011), has been suggested by the same authors (Holtkamp and Brudzinski, 2014) to reflect the existence of a low-coupling zone that could have contributed to stop the propagation of the Mw 9.2 1964 earthquake, although it is unclear whether this low-coupling zone extends sufficiently down-dip to effectively act as a barrier. We argue that transient A10 is possibly located in the transition zone between the locked asperity centered on Kodiak Island and the low-coupling zone extending between it and the Prince William Sound asperity. An instance of a (possibly repeating) SSE triggering seismicity in such a transition between strong and low coupling was found by Vallee et al. (2013) in another subduction context. A1 and A10 are located in creeping zones, while A3 and A8 are in locked zones, according to Freymueller et al. (2013). This shows that transient deformation generating seismicity is not a-priori constrained by large-scale coupling as inferred from GPS measurements. The total seismic moments of these swarms are of the order of 10^{17} Nm, i.e., M_w = 5.3 (Tables 1–2). This is of the same order as for example the 2007 Boso swarm evaluated to 3 × 10^{17} Nm (Ozawa et al., 2007). This Boso swarm was associated with a slow slip event, which released a total moment of 1.09 × 10^{18} Nm as given by inversion of GPS displacements. Unfortunately, there is no available GPS data to further document our transients, but the comparison to for example the 2007

Fig. 4. Location of the swarms, related focal mechanisms and magnitude versus time. The iso-depth contours are displayed with increments of 20 km starting from the trench in dashed red line (Hayes et al., 2012). We show in the magnitude versus time plot all the earthquakes in the area delimited by the corresponding map, for the particular time interval of interest. The earthquakes belonging to the swarm are highlighted in red.
Boso swarm suggests that the total aseismic slip could be much larger than the total aseismic slip relaxed by these swarms. As our transients do not occur repeatedly, at least over the 10 year long period studied here, and do not contain any large earthquake, they contribute very little to the local large-scale relaxation over the time scale of years, and therefore do not affect the local seismic coupling. This is unlike other known aseismic deformation transients, whose sizes and periodic dynamics contribute significantly to the coupling (Rogers and Dragert, 2003; Larson et al., 2007; Kostoglodov et al., 2010).

Our detected transients are indicative of a non-periodic slow slip instability that can take place in an otherwise locked zone, as observed prior to the 2011 $M_w$ 9.0 Tohoku earthquake. We did not detect any transient associated with the SSE of 1996–2001 in southern Alaska (Ohta et al., 2006; Fu and Freymueller, 2013; Peterson and Christensen, 2009). This SSE is located at depths between 30 and 50 km, below a strong asperity resisting the subducting movement of the Pacific plate. We infer that no significant seismicity was triggered by this slow slip because of its depth, as is also for example the case of Cascadia SSEs (Delahaye et al., 2009; Vidale et al., 2011).

It is interesting to note that most swarms are located close to the termination of mega-thrust recent ruptures (Fig. 5). Swarms $A_1$ and $A_3$ are located between the 1996 and 2003 earthquake rupture zones, and the $A_2$ transient occurs where the 1986 and 1996 earthquake ruptures overlap, as shown by their aftershock areas. However, the 1957 $M_w$ 9.1 earthquake ruptured this zone. While the $A_1$ and $A_3$ transients can be thought as being markers of a structurally complex zone that inhibits the propagation of large ruptures, the nature of the zone surrounding $A_2$ is less clear, and could perhaps be an area of low stress resulting from the 1957 $M_w$ 9.1 earthquake, hence a dynamically rather than structurally controlled barrier.

5. Conclusion

We have further developed the methodology of Marsan et al. (2013a) that search for seismicity patterns related to aseismic transients, by measuring the significance of these anomalies through model comparison with a null-hypothesis of no transient changes in background rate. Tests have been conducted to show that the transients are best detected when the detection parameters, or scales, $L$ and $\tau$, are close to the actual scales of the anomaly. Further improvement of the method would be needed to better separate anomalously strong aftershock sequences which, in our approach, can indeed be seen as anomalies, from swarm activity unrelated to large mainshock occurrence, that would mark on-going slow, aseismic deformation processes. A preliminary distinction can be obtained by exploiting the NND method of Zaliapin and Ben-Zion (2013a).

The Aleutian subduction zone exhibits six clear seismicity anomalies related to slow deformation transients, for the 2004–2013 period, and for $m \geq 3.5$ earthquakes. Out of these six anomalies, five affect the subducting plate, at depths ranging from 20 to 40 km. They tend to occur in areas and depths characterized by an intermediate seismic coupling, although the correlation is weak and would require many more samples to be made significant. Also, three of these transients are found at the terminations of past large to giant ruptures, and could this be seen as dynamically or structurally-controlled barriers. Finally, only one transient ($A_1$) counts a $m \geq 6.5$ shock, with precursory activity that resembles a foreshock sequence. Other large earthquakes do not show...
statistically significant precursory swarm-like activity, at least at the \( m_c = 3.5 \) cut-off.

**Acknowledgements**

We would like to thank Michel Bouchon, Jean-Robert Grasso and Jean Vandemeulebroek for stimulating discussions on this topic, as well as the editor, an anonymous reviewer and Stephen Holtkamp for constructive remarks. The latter suggested the role of our detected deformation transients as barriers to large earthquake ruptures. This work was financially supported by the French ANR ASEISMIC and the EC REAKT projects.

**Appendix A. Tests and model limitations**

In this appendix, we test our model on synthetic catalogs to evaluate its limitations.

**A.1. Test 1: Dependence on detection parameters**

We construct a synthetic catalog of seismicity with the same features as the Aleutian seismic catalog (our study area in this paper) but uniformly distributed in space. This catalog consists of \( N = 3226 \) events with magnitude ranging between 2 and 5.9, in an area of \( 600 \times 600 \) km\(^2\) and during \( T = 3648 \) days (see Fig. S1 in supplementary material). Earthquake magnitudes follow a Gutenberg–Richter law with \( b = 1 \). The ETAS parameters inverted for this catalog are \( \alpha = 1.525, \beta = 1.135, c = 0.002 \) day, \( y = 2.450, L_0 = 0.100 \) km, \( k_0 = 0.014 \). We add to this catalog a background transient (called T1). This transient increases locally the background rate 1473 times, from \( 1.51 \times 10^{-6} \) earthquakes/day/km\(^2\) to \( 2.22 \times 10^{-3} \) earthquakes/day/km\(^2\). It starts at \( t = 3258 \) day, lasts for 5 days, affects a disk of 50 km radius and generates 71 events, either background or triggered by new earthquakes.

We first apply our method with detection scales \( \tau = 5 \) days and \( \mathcal{L} = 50 \) km, i.e., the same scales as the transient (see Fig. A.1).

We find that this transient is effectively detected with a significance level greater than 99% for these detection parameters. Our discretization in space and time happens to divide this 5-day, 50-km radius transient into neighboring cells instead of just one.

We also tested if the transient is detected with other detection parameters. We used \( \mathcal{L} = 30, 50 \) and 100 km, and \( \tau = 1, 5, 10 \) and 30 days. All 12 combinations of these detection parameters gave a >99% significance for at least one of the three cells affected by the transient.

We then add a second transient T2, less intense than T1. It has a total duration of 2 days, but its maximum activity is focused on a one-day period. It affects a 50 km-radius disk, and corresponds to an increase of the background rate by a factor of 400. It generates 8 new earthquakes. It is located 30 km away from T1 and occurs about 4 days before it. We again check how the detection of T2 depends on detection parameters \( \mathcal{L} \) and \( \tau \). Fig. A.2 summarizes our results. The transient can be detected with a significance level greater than 70%, only when the detection parameters \( \mathcal{L} \) and \( \tau \) are similar to the size and duration of the transient. In the next section, we present a simple analytical approximation to explain this result.

**A.2. Test 2: Dependence of the log-likelihood \( \Delta \ell \) on detection parameters**

We consider a spatio-temporal cell \( i \), for which the background seismicity estimated by the null hypothesis is \( \mu_0^{(i)} \) and the time-varying background rate is \( \mu_1^{(i)} \). Eq. (11) can be rewritten as

\[
\Delta \ell^{(i)} = \ell_1^{(i)} - \ell_0^{(i)} = -\left(\mu_1^{(i)} - \mu_0^{(i)}\right) \mathcal{L}^2 + \sum_{j=1}^{N} \ln\left(\frac{\mu_1^{(i)} + v_j}{\mu_0^{(i)} + v_j}\right)
\]

where \( N \) is the number of earthquakes inside cell \( i \).

We assume that during a transient the background rate increases up to \( \mu_1^{(i)} \gg \mu_0^{(i)} \). We also consider that the interaction terms are small, so that \( v_j \ll \mu_0^{(i)} \) for all earthquakes \( j \). We further assume that during the transient most earthquakes are background events, so that \( (\mu_1^{(i)} - \mu_0^{(i)}) \mathcal{L}^2 \approx N \).

We can thus approximate the gain in likelihood as

\[
\Delta \ell^{(i)} \approx -N + N \ln x
\]

where \( x = \mu_1^{(i)} / \mu_0^{(i)} \).

We denote by \( \mathcal{L}^* \) and \( \tau^* \) the size and duration of the transient, and by \( \Theta^* = [\mathcal{L}^*, \tau^*] \) the corresponding set of scales. We analyze how the gain in likelihood depends on the departure of the detection parameters \( \Theta = [\mathcal{L}, \tau] \) from \( \Theta^* \). To do so, we let one parameter vary while the other is kept constant.

If \( \Theta \ll \Theta^* \), i.e., we observe the seismicity over cells much smaller than the scale of the transient, the number of earthquakes \( N \) in the cell grows linearly with \( \Theta \), while \( x \) does not change. Since \( x \gg 1 \), we find that \( \Delta \ell^{(i)} \) also grows linearly: \( \Delta \ell^{(i)} \sim \Theta \). For
\( \Theta \gg \Theta^* \), the number \( N \) increases very slowly with \( \Theta \) since we only pick extra earthquakes outside the burst of activity caused by the transient, while the estimated gain in background rate \( x \) decreases as \( x \sim \frac{1}{\Theta} \). This causes \( \Delta N^{(i)} \) to decrease: \( \Delta N^{(i)} \sim -N \ln \Theta \).

This simple argument shows that the likelihood gain is effectively maximized when \( \Theta = \Theta^* \); a transient is more likely to be detected if the detection parameters are tuned to its characteristic scales (its significance is maximum). We test this behavior and the existence of the two end-member regimes using the synthetic catalog containing transient T2 (see Table S13). This table confirms the strong correlation between the maximum likelihood gain and the detection parameters when they are close to the transient scales.

### A.3. Test 3: Aftershock sequences seen as transients

The interaction model (ETAS) we use is based on ensemble-averaged laws. Departure from these laws is expected in reality. For example, the number of aftershocks directly triggered by a mainshock of magnitude \( m \) is, according to the model, a Poisson law, identical for all the mainshocks with magnitude \( m \); fluctuations in these numbers are not predicted by the model outside the natural dispersion of Poisson random deviates. A mainshock triggering many more aftershocks than predicted by the model will then lead to the detection of a transient with our approach.

While it could be argued that such a vigorous aftershock sequence is indeed the signature of an actual transient, for example corresponding to strong afterslip, it is difficult with our approach to evaluate its reality since we do not account for possible, natural departures from the ensemble-averaged laws. We therefore test whether a very vigorous aftershock sequence can cause our method to detect a transient. To do so, we focus on the biggest mainshock of the synthetic catalog, of magnitude \( m = 6 \). It triggers an aftershock sequence of about 300 aftershocks. We artificially decrease the magnitude of the mainshock from 6 to 3 in steps of 0.5, and apply our method with detection parameters \( \ell = 30 \) km and \( \tau = 1 \) day.

Fig. A.3 shows the probability that the sequence is identified as a transient. For a decrease of 1.5 of magnitude, the probability to find a transient exceeds 90%. In theory, this aftershock sequence is too strong by a factor \( e^{1.5} \approx 10 \) for a mainshock of magnitude \( m = 6 - 1.5 = 4.5 \). Departures from the mean productivity law by a factor of 10 have been observed (Tahir and Grasso, 2014) in real catalogs. We thus expect the detection of transients embedded in aftershock sequences. Section 3.3 describes a possible way to discriminate these from transients not related to highly productive mainshocks.

### Appendix B. Supplementary material

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.epsl.2014.12.012.

### References


