

TECHNICAL REPORT

Open Access



Non-stationary ETAS to model earthquake occurrences affected by episodic aseismic transients

Sasi Kattamanchi^{1,2*} , Ram Krishna Tiwari^{1,2} and Durbha Sai Ramesh³

Abstract

We present a non-stationary epidemic-type aftershock sequence (ETAS) model in which the usual assumption of stationary background rate is relaxed. Such a model could be used for modeling seismic sequences affected by aseismic transients such as fluid/magma intrusion, slow slip earthquakes (SSEs), etc. The non-stationary background rate is expressed as a linear combination of B-splines, and a method is proposed that allows for simultaneous estimation of background rate as well as other ETAS model parameters. We also present an extension to this non-stationary ETAS model where an adaptive roughness penalty function is used and consequently provides better estimates of rapidly varying background rate functions. The performance of the proposed methods is demonstrated on synthetic catalogs and an application to detect earthquake swarms (possibly associated with SSEs) in Hikurangi margin (North Island, New Zealand) is presented.

Keywords: Non-stationary ETAS, Hikurangi, Slow slip, Aseismic transients

Introduction

Earthquakes are generated by a complex system—the Earth. Lithospheric plates gliding past one another causes stress buildup in the crustal rocks which is released in short, sudden bursts causing earthquakes. While this plate motion is the major source of earthquake activity, several other factors like the heterogeneity of the earth's crust, the local stress conditions in the seismogenic area, fluid content, the disposition of existing faults and their stress histories, etc., also determine the size, location and time of occurrence of an earthquake. Since it is difficult to represent such a complex system physically, stochastic models are increasingly being used to model earthquake occurrences (Console et al. 2010; Helmstetter and Sornette 2002; Kagan and Knopoff 1981; Marsan and Lengline 2008; Ogata 1988, 1999). Well-established empirical laws like the Gutenberg-Richter law and the Omori's law facilitate the formulation of stochastic

models of earthquake occurrences such as the epidemic-type aftershock sequence (ETAS) model (Console 2003; Helmstetter 2003; Ogata 1988, 1998).

The ETAS model (Ogata 1988, 1998) is currently the most popular model to describe seismicity in a region and to test various hypotheses related to it. It is based on the premise that every earthquake has a magnitude-dependent ability to trigger aftershocks. The model considers that the earthquake sequence comprises aftershocks (events triggered by other earthquakes) and background earthquakes (events that occur independent of other earthquakes). The aftershocks are triggered because of internal stress adjustments in the seismogenic system initiated by the occurrence of an earthquake, while the background earthquakes are generally caused by forces related to plate tectonics. In addition, aseismic transient forces related to fluid/magma intrusion, slow slip etc., can also trigger (swarms of) earthquakes. Because such earthquakes do not follow typical mainshock–aftershock patterns, they are included in the model as background events. To analyze earthquake sequences that include such swarms which are short-lived, the usual assumption of stationary background activity needs to be relaxed.

*Correspondence: sasikattamanchi@ngri.res.in

¹ Academy of Scientific and Innovative Research (AcSIR), CSIR-National Geophysical Research Institute (CSIR-NGRI) Campus, Uppal Road, Hyderabad 500007, India

Full list of author information is available at the end of the article

This leads to an ETAS model with non-stationary background rate, referred to as non-stationary ETAS model in this study.

The traditional epidemic-type aftershock sequence (ETAS) model (Ogata 1988, 1998) assumes the background activity to be stationary and hence a single (constant) parameter (μ_0) is enough to represent it. On the contrary, a non-stationary background rate warrants representation by a continuous function $\mu(t)$ which needs to be estimated from the observed data. Some previous studies (Hainzl and Ogata 2005; Lombardi et al. 2006, 2010; Marsan et al. 2013a; Reverso et al. 2015) have presented analysis of earthquake occurrences (sequences) using ETAS models with non-stationary background rate. Initial studies such as Hainzl and Ogata (2005) and Lombardi et al. (2006, 2010) approximated the non-stationary background rate by fitting stationary ETAS model to data in moving windows. Later, Marsan et al. (2013a) adapted the iterative procedure of Zhuang et al. (2002) to the case of non-stationary ETAS model. However, these methods do not always produce a smooth solution and often result in an estimate of $\mu(t)$ that is wiggly, indicating overfit to the data. Marsan et al. (2013) described a procedure based on ETAS model to detect earthquake swarms triggered by aseismic transients. This was further extended in Reverso et al. (2015). However, the results from this method are sensitive to the choice of width of regular grid used to represent background rate. In a different approach, Llenos and McGuire (2011) estimated the background rate by modeling the aftershock activity using ETAS, and subtracting this from the total observed seismicity rate. An assumption of zero background rate is made in estimating the aftershock activity by fitting ETAS model to the complete sequence of earthquakes, and thus the estimate of $\mu(t)$ obtained by this method could be biased.

Recently, Kumazawa and Ogata (2014) adapted the hierarchical modeling approach of Ogata (2004, 2011) to non-stationary ETAS model. In this approach, the background rate function $\mu(t)$ is expressed as a piece-wise linear function made up of linear pieces between each pair of consecutive earthquake occurrence times. A two-step iterative procedure involving penalized maximum likelihood estimation (penalized MLE) and Type-II maximum likelihood method (Type-II MLE) was proposed for estimation of the piece-wise linear $\mu(t)$. In their method, however, only $\mu(t)$ is estimated while employing a priori ETAS model parameters. In their study, the ETAS model parameters were fixed to those obtained by fitting a stationary ETAS model to the earthquake catalog corresponding to a wider region around the study area. Often, such pre-determination of model parameters may

not be feasible because the seismicity is localized (e.g., reservoir associated seismicity) or the model parameters are spatially inhomogeneous. Aside, owing to the strong dependence between $\mu(t)$ and other ETAS model parameters (Harte 2013; Marsan et al. 2013a), input of erroneous model parameters to the algorithm would result in biased estimates of the background rate. Therefore, a method that can simultaneously estimate both, the model parameters and the background rate, becomes desirable. Our study attempts to address this issue.

The method proposed in this study is largely based on Kumazawa and Ogata (2014). We express the non-stationary background rate function $\mu(t)$ as a spline function. Instead of using all earthquake occurrence times as knots to represent this spline function, we express it as a linear combination of finite number of basis splines (B-splines), akin to the P-splines method of Eilers and Marx (1996). Additionally, we employ the L-curve procedure (Frasso and Eilers 2015) to choose the optimal smoothing parameter in place of Type-II MLE. The penalized MLE method is then used to simultaneously invert for both the background rate $\mu(t)$ and other ETAS model parameters related to aftershock activity. Such simultaneous estimation of all unknowns in the model produces more reliable results.

Further, except for the method of Kumazawa and Ogata (2014), none of the above cited methods employ an explicit roughness penalty function to constrain the wiggleness of the estimated $\mu(t)$, and therefore, could lead to estimates that overfit the data. Even the method of Kumazawa and Ogata (2014) uses a global smoothness parameter to weigh the roughness penalty function in the expression for penalized log-likelihood. Thus, these methods are less effective to estimate the background rate function in situations where significant non-uniformity in the roughness of $\mu(t)$ exists. This situation often arises while analyzing earthquake sequences of long duration that are affected by occasional aseismic transients. To model such sequences, we therefore extend the method of Kumazawa and Ogata (2014) by representing the smoothness parameter as yet another spline function.

We explore the performance of these proposed methods on synthetic catalogs simulated using two different types of background rate functions. Further, an application of these methods to earthquake catalog from Gisborne area near Hikurangi subduction zone (New Zealand) is presented. The estimated background rate exhibits a few peaks indicating brief periods of increased seismicity. Examination of continuous global positioning system (cGPS) data recorded at nearby stations suggests that these peaks coincide with periods of reversals in recorded displacements that indicate slow slip.

The ETAS model

The epidemic-type aftershock sequence (ETAS) model was introduced by Ogata (1988) and, since then, is widely used in the analysis of earthquake sequences. It is a point process model which assumes that the earthquake sequence is made up of aftershocks and background events. Aftershocks include those events that are triggered by other earthquakes while background events are those that occur independent of other earthquakes. So, according to the model, the sequence of observed earthquakes is generated by a non-homogeneous Poisson process with rate $\lambda(t)$ expressed as

$$\lambda(t) = \mu(t) + \nu(t) \quad (1)$$

where $\mu(t)$ and $\nu(t)$ are the rates of background and aftershock events, respectively.

Aftershock rate can be expressed in parametric form using well-known empirical characteristics of aftershocks—(a) the rate of aftershock activity triggered by an earthquake decays according to the Omori-Utsu law (Utsu et al. 1995) and (b) the total number of aftershocks triggered by an earthquake depends exponentially on its magnitude. Thus, the rate of aftershock activity triggered by an earthquake i occurring at time t_i with magnitude M_i is

$$\xi_i(t) = Ke^{\alpha(M_i - M_0)} (t - t_i + c)^{-p} \quad (2)$$

where $\{K, \alpha, c, p\}$ are unknown parameters and M_0 is the cutoff magnitude of the catalog. Note that only the earthquakes with magnitudes above the cutoff magnitude are considered for analysis using the ETAS model.

The total aftershock rate at any time t can thus be obtained as

$$\nu(t) = \sum_{i:t_i < t} Ke^{\alpha(M_i - M_0)} (t - t_i + c)^{-p} \quad (3)$$

Background rate is generally considered to be constant $\mu(t) = \mu_0$ in time. This is because, for the short time spans of observed catalogs, the effect of long-term plate tectonic forces, which are primarily responsible for background activity, can be considered to be constant. However, when aseismic transient forcings associated with slow slip earthquakes or fluid intrusion, etc., are believed to affect the observed earthquake sequence, a time-varying background rate function $\mu(t)$ is required for better modeling of data. A non-stationary ETAS model with such a time-varying background rate is considered in this study. We assume that the other model parameters are stationary over the observed time period.

Unlike aftershock rate, the background rate function $\mu(t)$ cannot be expressed in a general parametric form. Various authors have employed different ways to represent $\mu(t)$ and adopted different procedures for inversion/

estimation of the unknowns as enunciated in the earlier section. For example, Kumazawa and Ogata (2014) represented the background rate as piece-wise linear function and performed estimation using an iterative procedure involving penalized maximum likelihood and Type-II maximum likelihood method. On the other hand, Hainzl et al. (2013) and Marsan et al. (2013a) used a different iterative procedure where they (a) begin with a homogeneous ETAS model and estimate its model parameters; (b) compute the probabilities of each event to be a background event (given by $\mu(t_i)/\lambda(t_i)$); (c) estimate the background rate by smoothing these probabilities; (d) use this estimated $\mu(t)$ as the new background rate and re-estimate the other model parameters. The steps (b)–(d) are repeated till convergence to obtain final estimates.

In the present study, we express the background rate $\mu(t)$ as a spline function. Unlike Kumazawa and Ogata (2014) where all the earthquake occurrence times are used as internal knots of the spline function, we employ only a fixed number M of B-spline basis functions as in P-splines methodology (Eilers and Marx 1996). That is, $\mu(t)$ is represented by

$$\mu(t) = \sum_{i=1}^M \phi_i B_i(t, d, \kappa_t) \quad (4)$$

where M is the total number of B-splines, ϕ_i are spline coefficients, $B_i(t, d, \kappa_t)$, $i = 1, 2, 3, \dots, M$, are the B-spline basis functions of degree d computed over the knot vector κ_t . B-spline basis functions for any given knot vector and spline degree can be computed using the deBoor's algorithm (De Boor 1978). The degree and knot vector of B-splines determine the flexibility of the spline function in fitting the observations. A spline function of degree d would have continuous derivatives of up to $d - 1$. So, the larger the degree of spline function, the higher would be the smoothness. On the other hand, the knot vector determines the local flexibility of the spline function. The closer the knots, the more flexible would be the spline function in fitting the observations. In this study, we employ quantile spaced knots (Ruppert et al. 2003). That is, knots are chosen such that any two consecutive knots contain the equal number of events. This kind of knot placement is better suited for modeling any sudden surges in background activity compared to uniform knot placement.

Thus, the non-stationary ETAS model can be described as a non-homogeneous Poisson process with rate

$$\lambda(t) = \sum_{i=1}^M \phi_i B_i(t, d, \kappa_t) + \sum_{i:t_i < t} Ke^{\alpha(M_i - M_0)} (t - t_i + c)^{-p} \quad (5)$$

The unknowns in this model are $\Phi = \{\phi_i, i = 1, 2, 3, \dots, M\}$ corresponding to background activity, and other model parameters $\Theta = \{K, \alpha, c, p\}$ related to aftershock activity. Estimating these unknowns and plugging them in the above expression would give us the non-stationary ETAS model that best describes the observed earthquake sequence.

Estimation

In the simpler case of homogeneous ETAS model, where $\mu(t) = \mu_0$, the unknowns are only $\{\mu_0, K, \alpha, c, p\}$. They can be estimated using maximum likelihood estimation (MLE) method by maximizing the log-likelihood function $\log L$ (Ogata 1998)

$$\log L = \sum_{\{i:S < t_i < T\}} \log \lambda(t_i) - \int_S^T \lambda(t) dt \quad (6)$$

where $[S, T]$ is the time domain containing the observations.

However, simple maximum likelihood method does not provide good estimates in case of a non-stationary ETAS model. The MLE estimates would be such that the background rate overfits the data. To avoid this, a roughness penalty has to be applied to constrain the wiggleness of the estimated $\mu(t)$. Thus, the model unknowns Φ and Θ are estimated by penalized maximum likelihood estimation (penalized MLE) method. Estimated $\hat{\Phi}$ and $\hat{\Theta}$ are those that maximize the penalized log-likelihood objective (Kumazawa and Ogata 2014)

$$R(\Phi, \Theta, \tau) = \log L(\Phi, \Theta) - \tau \times Q(\Phi) \quad (7)$$

where τ is the regularization or smoothing parameter and $Q(\Phi)$ is the roughness penalty function. The roughness penalty is generally considered in the form of integrated squared m th-order derivative of the desired function

$$Q(\Phi) = \int_S^T [\mu^{(m)}(t)]^2 dt \quad (8)$$

When $\mu(t)$ is expressed in terms of B-spline basis, the penalty function can be conveniently expressed as

$$Q(\Phi) = \Phi' P \Phi \quad (9)$$

where the penalty matrix P is a symmetric matrix with elements

$$P_{ij} = \int_S^T B_i^{(m)}(t, d, \kappa_t) B_j^{(m)}(t, d, \kappa_t) dt \quad (10)$$

The smoothing parameter τ in Eq. (7) controls the relative contribution of goodness-of-fit criterion (here log-likelihood) and the roughness penalty function in determining the values of estimated parameters. Large τ values in a penalized MLE lead to over smoothed estimates of $\mu(t)$,

while a small τ results in a $\mu(t)$ which is under smoothed. It is therefore important to employ a τ that provides optimal smoothing.

Choosing optimal smoothness parameter

Consider the penalized log-likelihood objective function given in Eq. (7). The penalty function is used to regularize the inversion and the smoothness parameter τ plays the role of a regularization parameter. The purpose of the penalty function is to provide additional constraints to the inversion to impart stability.

As the penalty function depends only on Φ , it thus provides constraints for the parameters in Φ alone. Depending on the choice of the order m of roughness penalty, the penalty matrix P could be rank deficient. For a ridge penalty ($m = 0$), the penalty matrix would be full rank ($r = M$); for a penalty on first-order derivatives ($m = 1$), the penalty matrix would have a rank $M - 1$, and so on. In general, for a penalty on the m th-order derivative, the penalty matrix would be rank deficient by m and would have a rank of $r = M - m$. This rank deficiency implies that the penalty matrix cannot simultaneously constrain all the parameters ϕ_i in Φ ; only $r = M - m$ of them would be constrained. In brief, of the $M + 4$ number of unknown parameters (Φ of dimension M and Θ of dimension 4), there exist $M - m$ constrained parameters, say $\Phi_f = \{\phi_i, i = 1, 2, 3, \dots, M - m\}$, and $4 + m$ unconstrained parameters $\Theta \cup \Phi_d$ where $\Phi_d = \{\phi_i, i = M - m + 1, \dots, M\}$. Apart from these, the optimal smoothing parameter τ also has to be estimated. Kumazawa et al. (2016) assumed a priori Θ and were hence left only with constrained parameters Φ_f and the other unknowns $\eta = \{\Phi_d, \tau\}$. They employed the penalized MLE step to estimate Φ_f and the Type-II MLE to estimate η .

Type-II maximum likelihood estimation

Consider the expression for penalized log-likelihood given in Eq. (7). From the Bayesian perspective, applying a roughness penalty to the log-likelihood function is equivalent to putting a prior on the variables. To understand this better, let us exponentiate both sides of Eq. (7)

$$e^{R(\Phi, \Theta, \tau)} = e^{\log L(\Phi, \Theta)} \times e^{-\tau \times Q(\Phi, \tau)}. \quad (11)$$

In the above equation, $e^{-\tau \times Q(\Phi, \tau)}$ corresponds to prior, $e^{\log L(\Phi, \Theta)}$ is the likelihood and $e^{R(\Phi, \Theta, \tau)}$ is proportional to the posterior. Note that estimating the parameters that maximizes the penalized log-likelihood objective is same as finding the mode of the above posterior. Thus, penalized maximum likelihood estimation is equivalent to maximum a posteriori (MAP) estimation.

The prior in expression (11) is improper because of the mentioned rank deficiency of the penalty matrix. However,

the penalty matrix has rank $M - m$ and thus the prior is proper on the $M - m$ parameters Φ_f . So, we can normalize this part of the prior as $e^{-\tau \times Q(\Phi_f)} / \int e^{-\tau \times Q(\Phi_f)}$ to obtain a prior, which is a proper probability density function. Using this, the posterior $T(\Phi_f, \eta)$ can be written as

$$T(\Phi_f, \eta) = L(\Phi, \Theta) \times \frac{e^{-Q'(\Phi_f, \tau)}}{\int e^{-Q'(\Phi_f, \tau)}} \quad (12)$$

where $Q'(\Phi_f, \tau) = \tau \times Q(\Phi_f)$ is the roughness penalty scaled by the smoothing parameter τ .

Note that there is no prior on the parameters η . They are treated as hyperparameters and are estimated by maximizing the posterior that is marginalized over Φ_f as

$$\Lambda(\eta) = \int T(\Phi_f, \eta) d\Phi_f \quad (13)$$

This procedure of estimating hyperparameters by maximizing the marginalized likelihood is called the Type-II maximum likelihood procedure (Kumazawa and Ogata 2014; Ogata 2011). This procedure is also known as empirical Bayes (Bishop 2006). The Akaike Bayesian Information Criterion (ABIC) is equal to $-2 \times \max \log \Lambda(\eta) + \dim(\eta)$. Thus, estimation using Type-II MLE is equivalent to choosing a model that minimizes the Akaike Bayesian Information Criterion (ABIC).

Computing the marginalized posterior involves integration over $\Phi_f = \{\phi_i, i = 1, 2, 3, \dots, M - m\}$ which has large dimensionality. This integration is difficult to compute in practice; so, Laplace approximation is used to approximate the posterior by a Gaussian distribution and thereby simplify this integration. Using Laplace approximation, the logarithm of marginal likelihood can be written as (Ogata 2011)

$$\log \Lambda(\eta) = R(\hat{\Phi}_f | \eta) - \frac{1}{2} \log \det H_R(\hat{\Phi}_f | \eta) + \frac{1}{2} \log \det H_{Q'}(\hat{\Phi}_f | \eta) \quad (14)$$

where $H_R(\hat{\Phi}_f | \eta)$ and $H_{Q'}(\hat{\Phi}_f | \eta)$ denote the Hessians of penalized log-likelihood and roughness penalty, respectively, evaluated at $\hat{\Phi}_f$ corresponding to the peak of penalized log-likelihood function. These $\hat{\Phi}_f$ are generally unknown; in fact, it is our aim to estimate these. However, both Φ_f and η can be estimated by an iterative procedure consisting of two steps: (a) given η , Φ_f is estimated by penalized MLE and (b) using this estimate $\hat{\Phi}_f$, η is estimated using Type-II MLE by maximizing the approximated marginal likelihood objective given in Eq. (14). Φ_f and η are re-estimated in each iteration, until convergence. This iterative procedure was adopted by Kumazawa and Ogata (2014) to estimate Φ_f and $\eta = \{\Phi_d, \tau\}$ given a priori known Θ . Even

in the case when Θ are unknown, this iterative procedure can be used to estimate Φ_f (step (a)) and redefined $\eta = \{\Phi_d, \Theta, \tau\}$ (step (b)) as was done by Ogata (2011) in the context of spatially inhomogeneous ETAS. However, in practice, the results obtained using such an approach remain unsatisfactory, because the strongly dependent parameters Φ and Θ are estimated in two different steps. Thus, we propose to use the L-curve procedure of Frasso and Eilers (2015) to choose the optimal τ and then use penalized MLE to simultaneously estimate all model unknowns Θ and Φ .

L-curve method

It is well known that a small smoothing parameter τ results in a $\mu(t)$ which is wiggly (under smoothed), while a large τ gives an estimate of $\mu(t)$ that is over smoothed. Therefore, a good strategy is to perform penalized MLE over a range of τ , from small to large values, and then pick the optimal τ . When the roughness penalty values corresponding to the estimates of Φ and Θ at each τ are plotted against the respective negative log-likelihood values, an approximate L-shaped curve is obtained. The vertical part of the curve is associated with τ values that provide under smoothed solutions and the horizontal part with over smoothed solutions. Hence, the τ value corresponding to the corner of the L-curve can be chosen as the optimal smoothness parameter $\hat{\tau}$ (Frasso and Eilers 2015). The estimates from penalized MLE corresponding to this $\hat{\tau}$ are the optimal estimates of $\hat{\Phi}$ and $\hat{\Theta}$. This procedure is similar to the L-curve method for choosing the optimal regularization parameter in ill-posed inverse problems (Hansen 1999; Sen and Stoffa 2013).

Adaptive roughness penalty function

The above-described combination of penalized MLE and L-curve method may work well to effectively describe earthquake occurrences in most cases. However, in situations where the background rate function has significant non-uniform roughness over its time domain, use of a global smoothness parameter as indicated in the proposed method could be insufficient to model the earthquake sequence effectively. Quantile knots can take care of such variable smoothness to some extent. But, this is not always sufficient. To obtain better results in such cases, an adaptive penalty function is needed that allows local variations in roughness. In this study, such an adaptive penalty function is obtained by expressing the smoothness parameter as another spline function (Baladandayuthapani et al. 2005) defined over $M_\tau (<< M)$ B-spline functions as

$$\tau(t) = \sum_{i=1}^{M_\tau} \tau_i B_i(t, \kappa_\tau, d_\tau) \quad (15)$$

where τ_i , d_τ and κ_τ are the spline coefficients, degree and sub-knots corresponding to the smoothness parameter function $\tau(t)$. A small subset of background rate knots (κ_τ) are chosen as sub-knots (κ_τ) for τ . Thus, the penalized log-likelihood function in Eq. (7) would now be

$$R(\Phi, \Theta, \tau) = \log L(\Phi, \Theta) - \int_S^T \tau(t) [\mu^{(m)}(t)]^2 dt \quad (16)$$

where $\tau(t)$ is given by Eq. (15).

Since there are now more than one smoothness parameters (the spline coefficients τ_i), the L-curve methodology cannot be applied. However, the iterative procedure involving penalized MLE and Type-II MLE procedure (Kumazawa and Ogata 2014) described previously can be used to estimate the spline coefficients $\{\tau_i, i = 1, 2, \dots, M_\tau\}$ by including all of them in the hyper parameters vector η . Note that this method involving Type-II MLE gives reliable results only in the case when Φ is estimated given known Θ . Since model parameters are not always known beforehand, estimates are first obtained using the penalized MLE and L-curve approach described above. As we shall see in the section on synthetic tests, these are reasonably close to the true values. Thus, for estimation using the adaptive penalty approach, we first estimate (a) $\hat{\Theta}_L$, $\hat{\Phi}_L$ and $\hat{\tau}_L$ using penalized MLE and L-curve approach. (b) Set $\Theta = \hat{\Theta}_L$ (c) with $\Phi = \hat{\Phi}_L$ and $\tau_i = \hat{\tau}_L$, estimate $\hat{\tau}(i)$ and $\hat{\Phi}_A$ using the iterative procedure involving Type-II MLE (d) on convergence of the iterative procedure, use the newly estimated $\hat{\tau}(t)$ and perform a final penalized MLE step to re-estimate both Φ and Θ . In the case of a rapidly varying background rate function, as demonstrated in the following section, this adaptive procedure provides estimates of background rate that are much better than those obtained using a global smoothness parameter.

Synthetic tests

In this study, the purpose of testing is to learn if the combination of penalized MLE and the L-curve procedure is capable of determining all the model unknowns with reasonable accuracy. In addition, it is necessary to check whether all the model unknowns, $\mu(t)$ and Θ , are identifiable. That is, can all the unknowns be determined uniquely from a given combination of model and data? Non-identifiability would arise, say, if certain amount of background activity can be described as aftershock activity and vice versa. It is therefore imperative to determine existence of non-identifiability and assess its impact on the results. In order to achieve this, we additionally perform estimation where we invert only for $\mu(t)$ (or equivalently Φ) while constraining Θ to their true values. Large discrepancy in the estimated background rates obtained from the proposed simultaneous inversion scheme and that through constraining Θ is a reasonable indicator

of non-identifiability. In such a case, the model parameters Θ and Φ cannot both be determined from the data and further constraints become necessary to resolve the model parameters. Additionally, results from the proposed methods are compared with those obtained from Type-II MLE approach described above. Various approaches tested using synthetic catalogs are listed in Table 1 for ease of reference.

Synthetic catalogs are simulated using non-stationary ETAS model with (a) a Gaussian-type background rate and (b) an Omori's law-type background rate (Kumazawa and Ogata 2013; Marsan et al. 2013a). Both the background rate functions have an expected value of 500 over the time period of simulation. For each of these background rate functions and with a typical set of model parameters $\Theta_{\text{true}} = \{K = 0.008 \text{ events/day}, \alpha = 2, c = 0.01 \text{ day}, p = 1.1\}$, 100 synthetic datasets are simulated over a time period of [0, 500] days. The synthetic catalogs are simulated as a branching process (Zhuang and Touati 2015). First, the background events are simulated using the given non-stationary $\mu(t)$ function by thinning method. Then, for each background event, aftershock sequences are simulated. Time sorted sequence of all these events combined together forms a simulated catalog. Magnitudes of the events are generated to follow Gutenberg-Richter law with b-value one and their values between 2 and 8, using the inverse transform method (Felzer et al. 2002; Zhuang and Touati 2015). The number of earthquakes in each synthetic catalog simulated using Gaussian-type background rate function ranges from 664 to 3597 among which 93% of the catalogs contain less than 1200 events. On the other hand, the number of earthquakes in the catalogs simulated using Omori's law-type background rate lies between 651 to 5142 with 92% of the catalogs contain less than 1200 events. The number of background events in the catalogs range between [453,556] and [427,554] for the Gaussian-type and Omori's law-type synthetic datasets, respectively.

For all the non-stationary ETAS models tested in this study, we employ 100 linear B-splines to represent background rate and a roughness penalty on the first-order derivatives ($m = 1$). For the adaptive penalty method, we express the smoothness parameter τ as a piece-wise constant spline function on 10 sub-knots. Wherever necessary, L-curves are computed using a range of log-spaced smoothing parameters $\{10^{-4}, 10^{-3.5}, \dots, 10^{7.5}, 10^8\}$. Since L-curve procedure involves performing penalized MLE for a large number of τ values, we did not estimate L-curves for all the 100 synthetic catalogs as that would be computationally expensive. Because all of 100 catalogs (of each type) are simulated using the same background rate function, optimal τ estimated for one typical catalog would be approximately valid for all of them. Thus, we estimated optimal τ by computing L-curve only for

Table 1 List of different types of models considered in the current study

Type	Methods	Model parameters	Model abbreviation
Non-stationary	Penalized MLE + L-curve	Known	NS_L
Non-stationary	Penalized MLE + L-curve	Estimated	NS_L_MP
Non-stationary	Penalized MLE + Type-II	Known	NS_TII
Non-stationary	Penalized MLE + Type-II	Estimated	NS_TII_MP
Non-stationary	Penalized MLE + Adaptive τ	Estimated	NS_Adapt

one catalog of each type (see Figs. 1a, 2a) and used this τ value as the approximate optimal τ for all the 100 synthetic catalogs of the corresponding type of background rate function. We provide this optimal τ value as the initial τ value even for the models employing Type-II MLE (NS_TII_MP and NS_TII). Note that, for all the algorithms, we provide $\Theta = \Theta_{true}$ and $\phi_i = 1 \forall i$ as the initial values to the algorithm except for the model with adaptive penalty. For the adaptive penalty model, we provide the final estimates of Φ , Θ and τ from NS_L_MP model as the initial values.

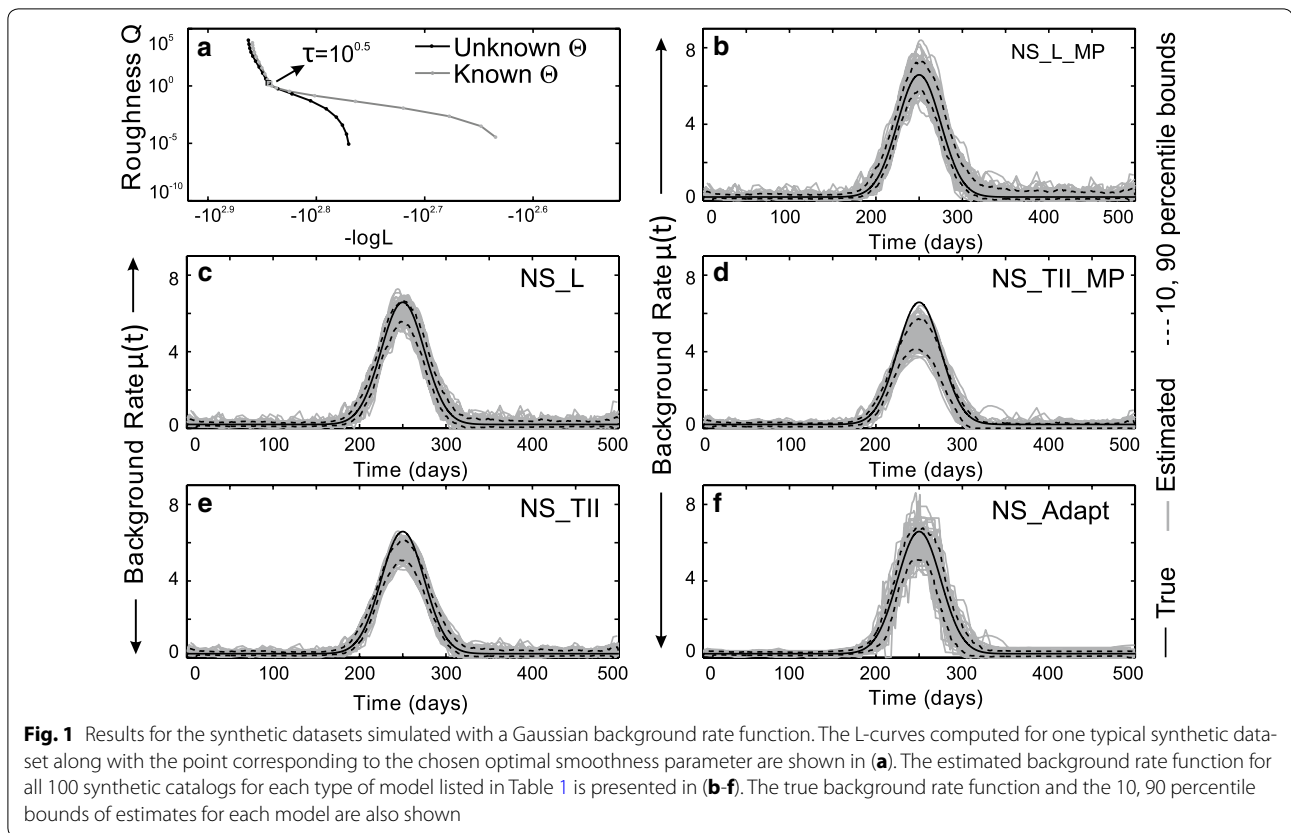
Figure 1b–f shows the background rates for all the 100 Gaussian-type synthetic catalogs, estimated using the non-stationary ETAS models listed in Table 1. It can be seen that for the simple case of Gaussian background rate, all the models are capable of providing good estimates of $\mu(t)$. Background rates estimated with models using a priori known model parameters (see Fig. 1c, e) and the models that allow simultaneous estimation along with $\mu(t)$ (see Fig. 1b, d) are nearly the same. This indicates that no significant non-identifiability exists and thus simultaneous estimation of both the model parameters and the background rate function $\mu(t)$ is feasible. Note that a few of the estimates obtained via adaptive penalty method (see Fig. 1f) show small block like distortions. This is caused by the usage of piece-wise constant τ function; over smoothing seems to occur in some windows. So, the adaptive penalty method is not very helpful in the case where the background rate is not rapidly varying (like the Gaussian background rate function).

In contrast, the results in the case of Omori's law-type background rate function indicate that the adaptive penalty method provides considerably better estimates of background rate. Estimates using the method employing L-curve approach with a global τ exhibit undesired wiggles (see Fig. 2b, c), whereas the estimates employing Type-II MLE approach seem to over smooth (see Fig. 2d, e), especially near the region with sudden jump. But using an adaptive penalty seems to damp the undesired wiggles (see Fig. 2f) and provide smoother estimates even while preserving the jump. So, the adaptive penalty method seems to be very helpful in this case of Omori's law-type background rate function whose roughness is much

more rapidly varying than the Gaussian background rate function.

Box plots of the estimated model parameters Θ using Models NS_L_MP, NS_TII_MP and NS_Adapt are shown in Figs. 3 and 4, respectively, for catalogs simulated using Gaussian-type and Omori's law-type background rates. Overall, both the proposed models NS_L_MP and NS_Adapt provide considerably better estimates than the model NS_TII_MP that employs Type-II MLE. Between these, the NS_Adapt model provides slightly better estimates as is evident by the closeness of the median values of the estimates to the true values, and lower interquartile ranges of the estimates. The difference is particularly prominent for the Omori's law-type synthetic catalogs. Observe that the estimates of parameter p obtained with the model NS_TII_MP are close to the true value and have a lower interquartile range as well, even while the estimates of other model parameters are bad. This is possibly due to the inability of Type-II MLE step in exploring values away from the initial value (= true value) provided.

The results reported above were computed using $M = 100$ B-splines to represent the background rate function $\mu(t)$ and $M_\tau = 10$ B-splines for the smoothness parameter $\tau(t)$. It is well known that the estimates obtained using the P-splines methodology are nearly independent of the number of splines used, provided they are sufficiently large in number to produce an overfitting estimate in the absence of any roughness penalty (e.g., Baladandayuthapani et al. 2005; Ruppert 2002). To test if this remains valid even for the methods described in the current study, we analyze a typical synthetic catalog of each type, using all possible combinations of $M = \{50, 100, 150, 200, 250, 300\}$ and $M_\tau = \{10, 20, 30, 40, 50, 60\}$ such that $M_\tau < M$. The estimates of the background rate function $\mu(t)$ thus obtained for each type of synthetic catalog using the proposed models M_L_MP and M_Adapt are shown in Fig. 5. It can be seen that all these combinations of M and M_τ yield similar estimates (see Fig. 5; Additional file 1: Tables S1, S2, S3 and S4 **in supporting information). Thus, these results suggest that the estimates obtained from the proposed methods are not very sensitive to the choice of the number of splines used to represent the background rate



and smoothness parameter. Although use of large number of splines M and M_τ may produce better results, the attendant computational cost increases steeply, especially with M_τ . Therefore, the choice of M and M_τ must be guided by the prior knowledge/expectation of the underlying background rate function $\mu(t)$. In any case, it would be prudent to examine the estimates using few other values of M and M_τ to confirm their stability.

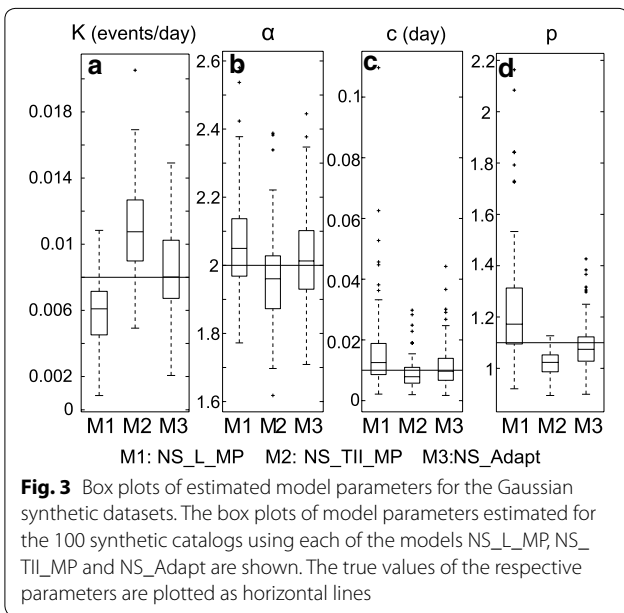
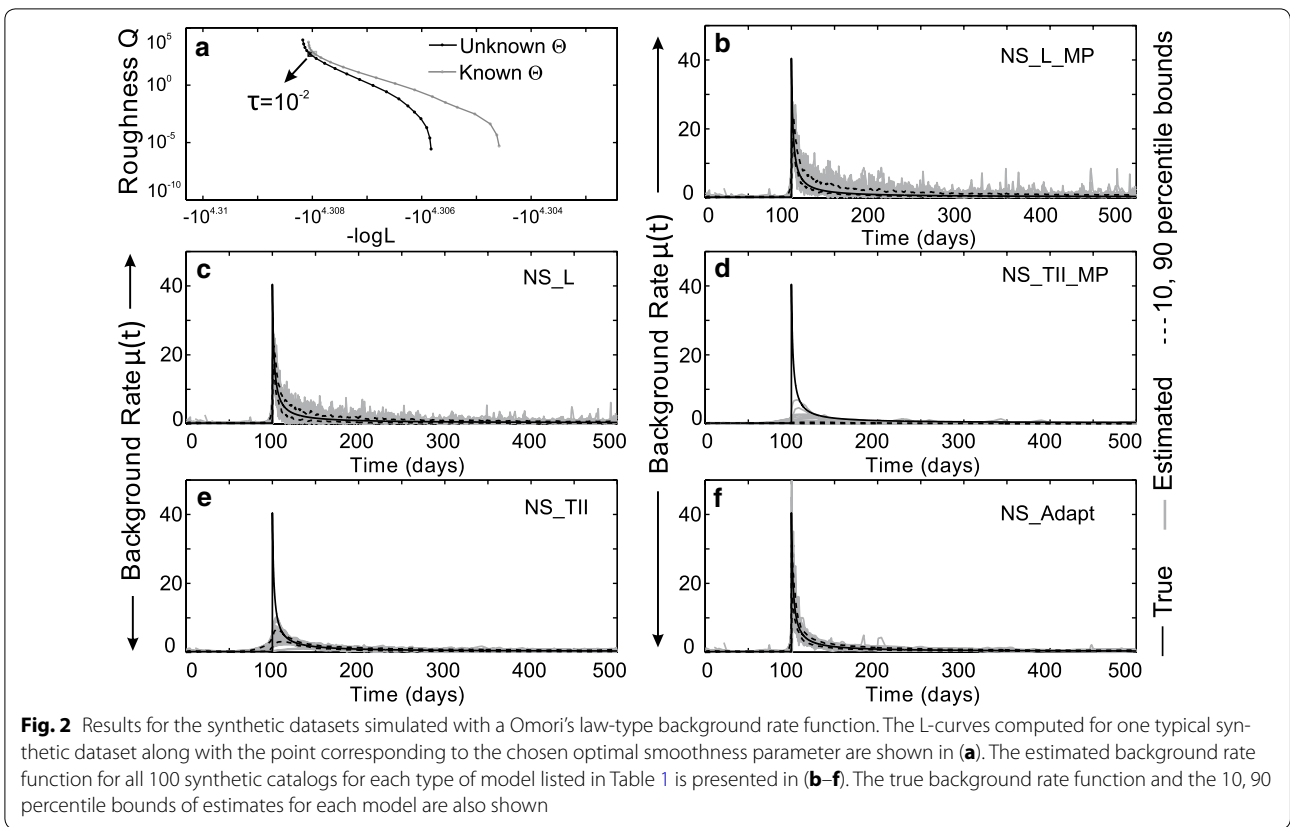
In the current study, we chose quantile based knot vector, for spline representation of $\mu(t)$, using all the earthquakes in the catalog. Therefore, more knots are inadvertently chosen even in the places where background rate is low but where aftershock activity is large. This can, in some cases, cause a small part of aftershock activity to be seen as background activity in the estimated model, owing to greater flexibility accorded by closely spaced knots. This is the cause of the occasional spurious bumps seen in the estimated background rate of synthetic catalogs. We observed that this issue chiefly arises when the analyzed catalogs contain large aftershock sequences and the background activity is such that a low smoothness parameter is needed. Using a low smoothness parameter τ further adds to the flexibility allowed by closely spaced knots and therefore can cause local overfitting which results in spurious bumps.

Adaptive penalties can alleviate this problem to a major extent. But, we speculate that devising and using a better knot selection strategy would also produce better and improved results.

Application to New Zealand data

We examine the GeoNet earthquake catalog of northern Hikurangi margin (New Zealand) and the cGPS data available for the region. The main objective is to demonstrate that the non-stationary ETAS model developed in this study is capable of identifying the transient increases in seismicity associated with slow slip earthquakes.

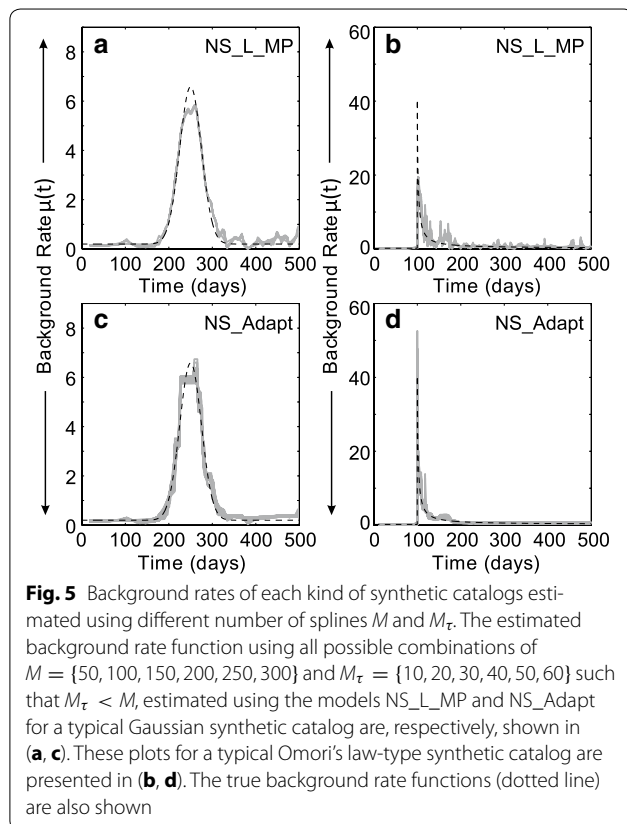
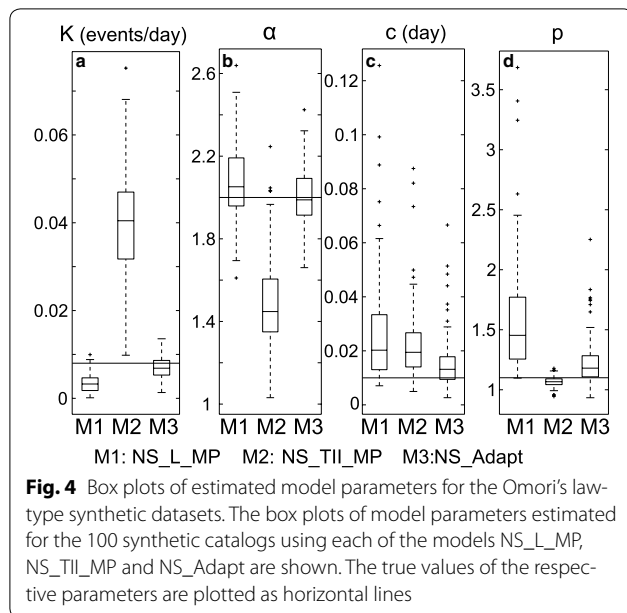
Slow slip earthquakes (SSEs) have been observed at many subduction zones around the world. They have been found to accompany with tremors in most cases, and thus together they have been called Episodic Tremor and Slip (Rogers and Dragert 2003; Schwartz and Rokosky 2007). Some of the slow slip earthquakes have also been found to trigger earthquake swarms (Delahaye et al. 2009; Hirose et al. 2014). On the other hand, some large earthquakes have also been believed to affect slow slip (Wallace et al. 2014; Zigone et al. 2012). It is necessary to understand such interactions between normal earthquakes and slow slip earthquakes in order to better appraise the seismic hazard potential of such regions.



At the Hikurangi margin, the Pacific plate is obliquely subducting beneath the Australian plate. Slow slip here occurs at a shallow depth similar to Boso peninsula, Japan. At both these subduction zones, the slow

slip earthquakes have been found to trigger earthquake swarms. Previous studies have identified earthquake swarms triggered by a few slow slip earthquakes in Hikurangi margin e.g., Gisborne 2004 SSE (Delahaye et al. 2009), Cape Turnagain 2011 SSE (Wallace et al. 2012), Puketiti 2010 SSE (Todd and Schwartz 2016). Here, we analyze the observed earthquakes near Gisborne (see Fig. 6), where multiple SSE have been documented, using the non-stationary ETAS model. From the estimated background rate, we would try to see if any anomalous seismicity is associated with slow slip earthquakes in the region.

The location algorithm and the reported magnitudes in GeoNet catalog changed since the beginning of 2012 (www.geonet.org.nz/data/supplementary/earthquake_location). Thus, for the analysis in this study we consider only the earthquake catalog since January 2012 until May 2017. The magnitude of completeness for the events in the target spatial window and the time period is estimated to be 1.9 M using maximum curvature method (Wiemer 2001). Thus, the non-stationary ETAS model is applied to 5007 events with magnitudes $M \geq 1.9$ (see Fig. 7a) and with depths less than 65 km, and whose epicenters are confined to the spatial window shown in Fig. 6. Given the large number of events in the



catalog and the frequent occurrence of SSEs that could affect the background seismicity in the region, we chose to use ($M =$) 150 linear B-splines to express the background rate function $\mu(t)$. A roughness penalty in terms

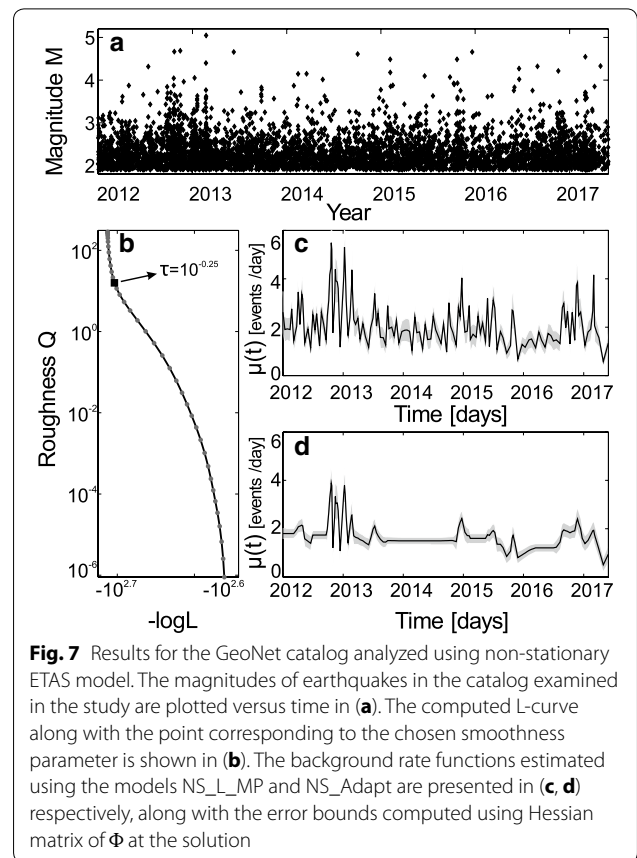
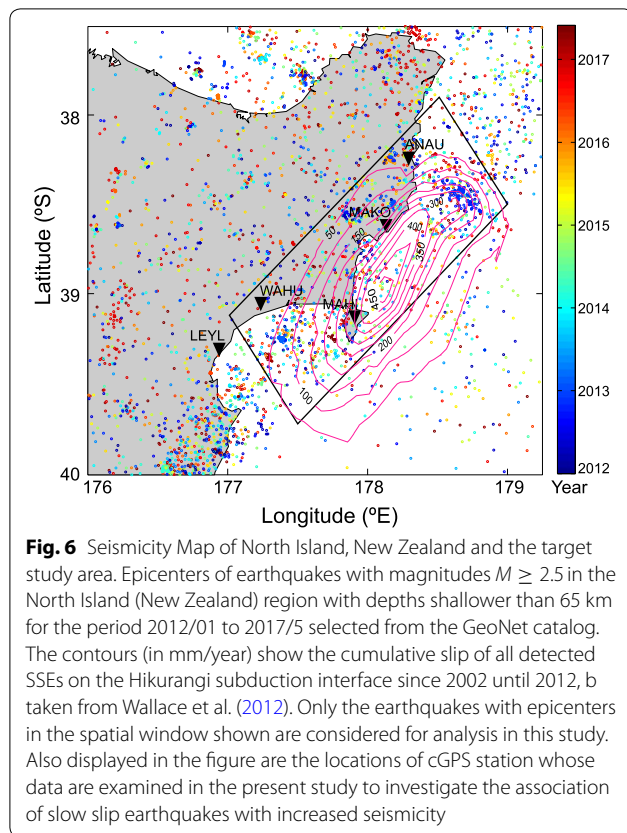
of first-order derivative ($m = 1$) is used. For modeling with adaptive penalty, the smoothness parameter $\tau(t)$ is expressed as a piece-wise constant spline made up of 30 B-splines.

The computed L-curve is shown in Fig. 7b, where the point corresponding to the chosen optimal smoothness parameters is marked. Estimated background rate function $\hat{\mu}(t)$, corresponding to this optimal $\hat{\tau}$, is wiggly (see Fig. 7c) and thus the adaptive penalty approach as described above is applied. The background rate (shown in Fig. 7d) thus estimated looks reasonably better than the one estimated without adaptive penalty. The estimated model parameters Θ corresponding to both the non-stationary ETAS models, with and without adaptive penalty, along with the standard errors are presented in Table 2. Note that the standard errors are computed as square root of diagonal elements of covariance matrix. The covariance matrix is estimated by taking inverse of the Hessian matrix at the solution. For goodness-of-fit tests, see supplementary material.

To see if slow slip earthquakes are associated with increased seismicity that manifests as peaks in the estimated background activity, we look at cGPS data recorded at the nearby cGPS stations LEYL, WAHU, MAHI, MAKO and ANAU (see Fig. 8). Slip caused by slow earthquakes is recorded by continuous GPS (cGPS) stations as reversals in the direction of slip over time periods ranging from a few days to years. Comparing the estimated background rate function with the east component of cGPS data recorded at the said cGPS stations, it is apparent that the peaks in estimated background rate are associated with deviations in cGPS east-west displacement time series indicative of slow slip earthquakes (see Fig. 8).

To check the sensitivity of the estimated background rate (shown in Fig. 8f) on the number of splines employed, we repeat the analysis described for the synthetic catalogs in the previous section, using different combinations of M and M_τ . The estimates of $\mu(t)$ thus obtained using the non-stationary ETAS model with adaptive penalty are shown in Fig. 8g and the corresponding model parameters in Table S5. These suggest that the considered combinations of M and M_τ produced similar estimates and reconfirm the presence of peaks in the background rate observed earlier (see Fig. 8).

It is pertinent to note that not all SSEs are associated with peaks in the background rate. For example, SSE in October 2014 do not seem to produce any pronounced increase in seismic activity, although there seems to be a faint peak in $\mu(t)$ (Fig. 8g) associated with this event. Lack of such recognizable increase in seismicity associated with this particular SSE was also observed by Todd and Schwartz (2016). They explain that as the SSE



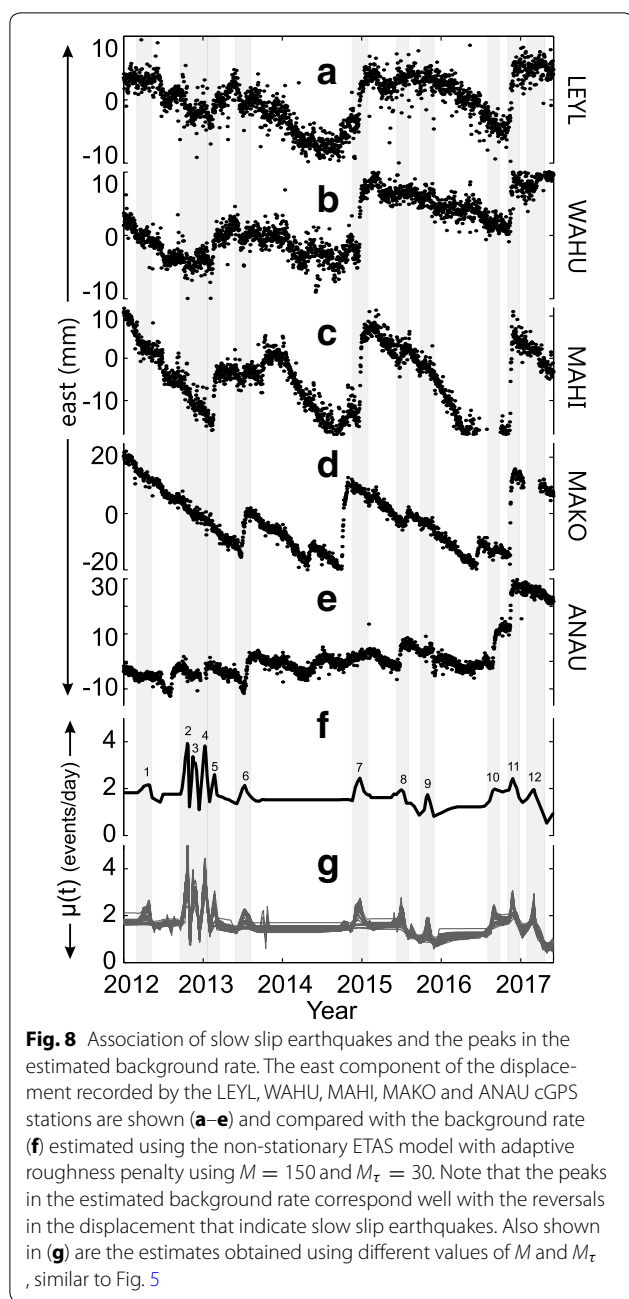
occurred far away from the shore, the associated increase in seismicity, if any, was not recorded by the onshore network. Employing a better catalog with a lower magnitude cutoff could possibly help identify more earthquake swarms associated with SSEs.

We plot the epicenters of the earthquakes that occurred during the time windows (shaded regions) corresponding to each of the peaks shown in Fig. 8. These scatter plots of earthquake locations (shown in Fig. 9) are visually enhanced adding color to individual plots based on smoothed histogram as described in Eilers and Goeman (2004). Such visual enhancement leads to better identification of spatial clusters of earthquakes. In each of the plots (Fig. 9), corresponding to peaks 2, 3, 4, 5, 7, 9, 11 and 12, distinct spatial clustering of earthquakes is visible. Approximate regions of slow slip available for a few SSEs (Koulali et al. 2017; Wallace and Eberhart-Phillips 2013; Wallace et al. 2017) are shown in the corresponding plots. It can be noticed that the seismicity clusters related to peaks 5 and 11 are located near the down dip edges of the corresponding slip patches (Fig. 9e, k). However, the earthquake cluster in Fig. 9k seems to be located within the slip region while that in Fig. 9e lies even outside the slip region. While the former lends support to the hypothesis that slow slip triggers earthquakes on

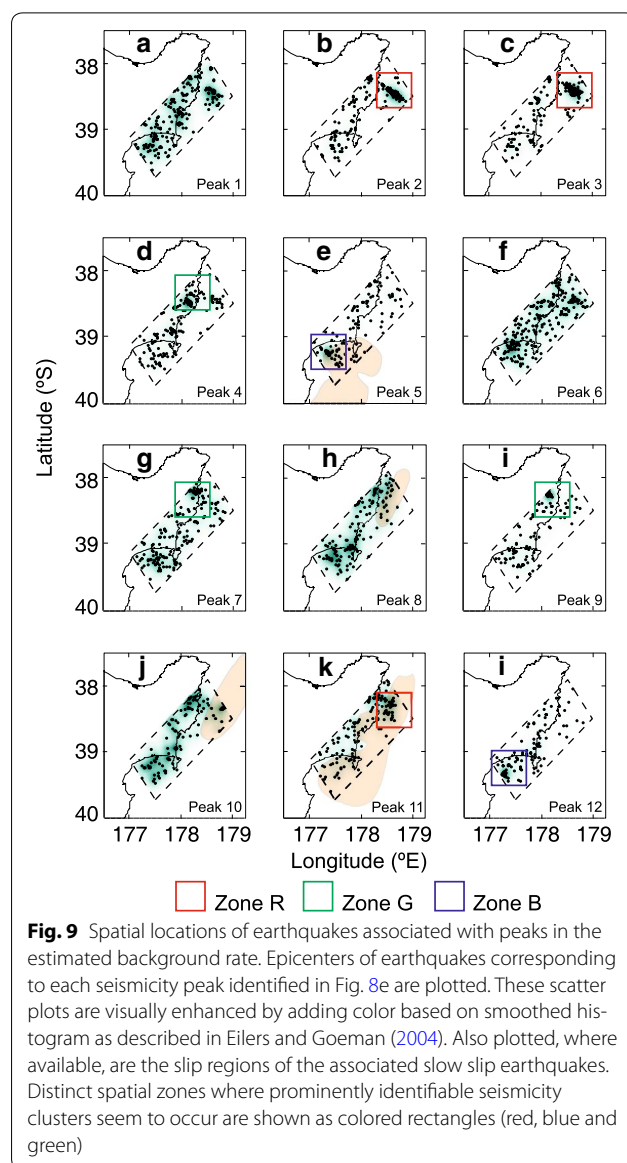
Table 2 Estimates of the ETAS model aftershock parameters ($\hat{\Theta}$) for the earthquake occurrences in Gisborne region for the period 2012/01–2015/05

$\hat{\Theta}$	NS_L_MP	NS_Adapt
K (events/day)	0.0067 ± 0.0018	0.0133 ± 0.0023
a	1.6137 ± 0.0713	1.5593 ± 0.0681
c (day)	0.0167 ± 0.0053	0.0116 ± 0.0033
p	1.3903 ± 0.1359	1.1516 ± 0.0750

locked asperities within the slip zone, the latter suggests triggering by increased static stress. Examination of the subplots in Fig. 9 suggests that the distinct spatial clusters of earthquakes associated with SSEs occupies three narrow spatial zones marked by colored rectangles (R, G and B). Spatial zone R (Fig. 9) coincides with a region where slow slip earthquakes have been detected in earlier studies (e.g., Koulali et al. 2017; see also contours in Fig. 6). Hence, it is possible that the earthquake swarms in this region are mainly caused by failure of locked asperities within the slip region. In contrast, earthquake clusters in spatial zone G are located inland, while slow slip



occurrence region is situated almost exclusively offshore (Todd and Schwartz 2016). Thus, the earthquake swarms falling in zone G, if triggered by SSEs, must have been caused by the associated increase in static stress. Previous studies on SSEs near Gisborne, that is spatial zone B, found that the associated swarms of earthquakes occur close to the down dip edge of the slip area (Bartlow et al. 2014; Delahaye et al. 2009). Thus, these swarms in spatial zone B are most likely triggered by static stress increase (Delahaye et al. 2009) akin to region G. Irrespective of



the triggering mechanisms, it is important to understand if these spatial zones are particularly susceptible to SSE-associated seismicity. Detailed modeling of slow slip events and estimation of associated Coulomb stress changes might shed more light in this direction. In addition, if epicentral locations too are modeled along with the earthquake occurrence times, using a non-stationary spatio-temporal ETAS model, spatial clusters of earthquakes associated with SSEs could be better identified.

Conclusions

In this study, we propose a P-splines-based non-stationary ETAS model and an estimation procedure involving (a) penalized maximum likelihood estimation and

(b) L-curve method for choosing optimal smoothness parameter. This procedure allows for simultaneous estimation of both background rate function and the other ETAS model parameters. Such a non-stationary ETAS model is useful in modeling earthquake sequences affected by time varying processes such as fluid/magma intrusion. For example, modeling the earthquake sequence associated with a particular swarm would help in understanding the time evolution of that swarm and could provide useful insights into the underlying causative process. We also present a non-stationary ETAS model that employs adaptive roughness penalty function. Such a model provides superior results when the background rate has significant non-uniform smoothness over its domain. This adaptive penalty method is particularly useful when analyzing long duration earthquake catalogs belonging to regions affected by occasional aseismic transients. The performance of both the proposed methods was demonstrated on synthetic datasets. An application to data from Hikurangi margin (New Zealand) is presented where the observed earthquake sequence near Gisborne is analyzed to find instances of increased background rate (earthquake swarms). These episodes of increased seismicity are then compared with cGPS data to better understand their association with slow slip earthquakes. The non-stationary ETAS model and the estimation procedures described in this study allow us to model earthquake activity affected by transient aseismic processes and thus allow us to obtain meaningful insights into these processes.

Additional file

Additional file 1. Goodness-of-fit tests for the modeled GeoNet catalog; Tables S1–S5 and Figure S1.

Abbreviations

ETAS: epidemic-type aftershock sequence; SSE: slow slip earthquakes; MLE: maximum likelihood estimation; MAP: maximum a posteriori estimation; ABIC: Akaike Bayesian Information Criterion; cGPS: continuous global positioning system.

Authors' contributions

SK developed the code, performed analysis, and drafted the manuscript. DSR participated in the design of the study. All authors participated in discussions and equally contributed to revising an earlier draft of the manuscript. All authors read and approved the final manuscript.

Author details

¹ Academy of Scientific and Innovative Research (AcSIR), CSIR-National Geophysical Research Institute (CSIR-NGRI) Campus, Uppal Road, Hyderabad 500007, India. ² CSIR-National Geophysical Research Institute, Uppal Road, Hyderabad 500007, India. ³ Indian Institute of Geomagnetism, New Panvel, Mumbai 410218, India.

Acknowledgements

We acknowledge the New Zealand GeoNet project and its sponsors EQC, GNS Science and LINZ, for providing the data used in this study. Dr. S. Das Sharma

is thanked for numerous discussions during this work. Shri Appala Raju is gratefully acknowledged for his technical help. We thank the two anonymous reviewers for their constructive comments. SK gratefully acknowledges funding from Council of Scientific and Industrial Research, New Delhi, India, through Shyama Prasad Mukherjee Fellowship.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

The data used in the study can be obtained from GeoNet Web site www.geonet.org.nz

Ethics approval, consent to participate and consent to publish

Not Applicable

Funding

This research was supported by funding from Council of Scientific and Industrial Research, New Delhi, India, through Shyama Prasad Mukherjee Fellowship (SPM-31/023(0178)/2013-EMR-I).

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 24 August 2017 Accepted: 26 October 2017

Published online: 14 November 2017

References

- Baladandayuthapani V, Mallick BK, Carroll RJ (2005) Spatially adaptive Bayesian penalized regression splines (P-splines). *J Comput Graph Stat* 14(2):378–394. <http://www.tandfonline.com/doi/abs/10.1198/106186005X47345>
- Bartlow NM, Wallace LM, Beavan RJ, Bannister S, Segall P (2014) Time-dependent modeling of slow slip events and associated seismicity and tremor at the Hikurangi subduction zone, New Zealand. *J Geophys Res Solid Earth* 119(1):734–753. <http://onlinelibrary.wiley.com/doi/10.1002/2013JB010609/abstract>
- Bishop CM (2006) Pattern recognition and machine learning. Springer. http://cds.cern.ch/record/998831/files/9780387310732_TOC.pdf
- Console R (2003) Refining earthquake clustering models. *J Geophys Res* 108(B10). <http://doi.wiley.com/10.1029/2002JB002130>
- Console R, Jackson DD, Kagan YY (2010) Using the ETAS model for catalog declustering and seismic background assessment. *Pure Appl Geophys* 167(6–7):819–830. <http://link.springer.com/10.1007/s00024-010-0065-5>
- De Boor C (1978) A practical guide to splines, vol 27. Springer, New York. https://www.researchgate.net/profile/Carl_De_Boor/publication/200744645_A_Practical_Guide_to_Spline/links/02e7e51700ff60945400000.pdf
- Delahaye E, Townend J, Reyners M, Rogers G (2009) Microseismicity but no tremor accompanying slow slip in the Hikurangi subduction zone, New Zealand. *Earth Planet Sci Lett* 277(1–2):21–28. <http://linkinghub.elsevier.com/retrieve/pii/S0012821X08006419>
- Eilers PHC, Goeman JJ (2004) Enhancing scatterplots with smoothed densities. *Bioinformatics* 20(5):623–628. <https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btg454>
- Eilers PHC, Marx BD (1996) Flexible smoothing with B-splines and penalties. *Stat Sci* 11(2):89–121. <http://projecteuclid.org/euclid.ss/1038425655>
- Felzer KR, Becker TW, Abercrombie RE, Ekström G, Rice JR (2002) Triggering of the 1999 M_w 7.1 Hector Mine earthquake by aftershocks of the 1992 M_w 7.3 Landers earthquake: TRIGGERING OF THE HECTOR MINE EARTHQUAKE. *J Geophys Res Solid Earth* 107(B9):ESE 6-1–ESE 6-13. <http://doi.wiley.com/10.1029/2001JB000911>
- Frasso G, Eilers PH (2015) L- and V-curves for optimal smoothing. *Stat Model* 15(1):91–111. <http://smj.sagepub.com/cgi/doi/10.1177/1471082X14549288>

- Hainzl S, Ogata Y (2005) Detecting fluid signals in seismicity data through statistical earthquake modeling. *J Geophys Res* 110(B5). <http://doi.wiley.com/10.1029/2004JB003247>
- Hainzl S, Zakharova O, Marsan D (2013) Impact of aseismic transients on the estimation of aftershock productivity parameters. *Bull Seismol Soc Am* 103(3):1723–1732. <http://www.bssaonline.org/content/103/3/1723>
- Hansen PC (1999) The L-curve and its use in the numerical treatment of inverse problems. IMM, Department of Mathematical Modelling, Technical University of Denmark. <http://www.sintef.no/globalassets/upload/ikt/9011/simoslo/vskoler/2005/notes/lcurve.pdf>
- Harte DS (2013) Bias in fitting the ETAS model: a case study based on New Zealand seismicity. *Geophys J Int* 192(1):390–412. <http://gji.oxfordjournals.org/cgi/doi/10.1093/gji/ggs026>
- Helmstetter A (2003) Is earthquake triggering driven by small earthquakes? *Phys Rev Lett* 91(5):058,501. <http://link.aps.org/doi/10.1103/PhysRevLett.91.058501>
- Helmstetter A, Sornette D (2002) Subcritical and supercritical regimes in epidemic models of earthquake aftershocks. *J Geophys Res Solid Earth* 107(B10):ESE 10-1–ESE 10-21. <http://onlinelibrary.wiley.com/doi/10.1029/2001JB001580/abstract>
- Hirose H, Matsuzawa T, Kimura T, Kimura H (2014) The Boso slow slip events in 2007 and 2011 as a driving process for the accompanying earthquake swarm. *Geophys Res Lett* 41(8):2778–2785. <http://onlinelibrary.wiley.com/doi/10.1002/2014GL059791/abstract>
- Kagan YY, Knopoff L (1981) Stochastic synthesis of earthquake catalogs. *J Geophys Res Solid Earth* 86(B4):2853–2862. <http://onlinelibrary.wiley.com/doi/10.1029/JB086iB04p02853/abstract>
- Koulali A, McClusky S, Wallace L, Allgeyer S, Tregoning P, D'Anastasio E, Benavente R (2017) Slow slip events and the 2016 te araroa mw 7.1 earthquake interaction: Northern hikurangi subduction, New Zealand. *Geophys Res Lett* 44(16):8336–8344. <https://doi.org/10.1002/2017GL074776>
- Kumazawa T, Ogata Y (2013) Quantitative description of induced seismic activity before and after the 2011 Tohoku-Oki earthquake by nonstationary ETAS models: NONSTATIONARY ETAS MODEL. *J Geophys Res Solid Earth* 118(12):6165–6182. <http://doi.wiley.com/10.1002/2013JB010259>
- Kumazawa T, Ogata Y (2014) Nonstationary ETAS models for nonstandard earthquakes. *Ann Appl Stat* 8(3):1825–1852. <http://projecteuclid.org/euclid.aas/1414091236>
- Kumazawa T, Ogata Y, Kimura K, Maeda K, Kobayashi A (2016) Background rates of swarm earthquakes that are synchronized with volumetric strain changes. *Earth Planet Sci Lett* 442:51–60. <http://linkinghub.elsevier.com/retrieve/pii/S0012821X16300747>
- Llenos AL, McGuire JJ (2011) Detecting aseismic strain transients from seismicity data. *J Geophys Res Solid Earth* 116(B6):B06,305. <http://onlinelibrary.wiley.com/doi/10.1029/2010JB007537/abstract>
- Lombardi AM, Marzocchi W, Selva J (2006) Exploring the evolution of a volcanic seismic swarm: the case of the 2000 Izu Islands swarm. *Geophys Res Lett* 33(7):L07,310. <http://onlinelibrary.wiley.com/doi/10.1029/2005GL025157/abstract>
- Lombardi AM, Cocco M, Marzocchi W (2010) On the increase of background seismicity rate during the 1997–1998 Umbria-Marche, Central Italy, sequence: apparent variation or fluid-driven triggering? *Bull Seismol Soc Am* 100(3):1138–1152. <http://www.bssaonline.org/cgi/doi/10.1785/0120090077>
- Marsan D, Lengline O (2008) Extending earthquakes' reach through cascading. *Science* 319(5866):1076–1079. <http://www.sciencemag.org/cgi/doi/10.1126/science.1148783>
- Marsan D, Prono E, Helmstetter A (2013a) Monitoring aseismic forcing in fault zones using earthquake time series. *Bull Seismol Soc Am* 103(1):169–179. <http://bssa.geoscienceworld.org/content/103/1/169>
- Marsan D, Reverso T, Helmstetter A, Enescu B (2013b) Slow slip and aseismic deformation episodes associated with the subducting Pacific plate off-shore Japan, revealed by changes in seismicity: ASEISMIC TRANSIENTS IN SUBDUCTION ZONE. *J Geophys Res Solid Earth* 118(9):4900–4909. <http://doi.wiley.com/10.1002/jgrb.50323>
- Ogata Y (1988) Statistical models for earthquake occurrences and residual analysis for point processes. *J Am Stat Assoc* 83(401):9–27
- Ogata Y (1998) Space-time point-process models for earthquake occurrences. *Ann Inst Stat Math* 50(2):379–402. <http://link.springer.com/article/10.1023/A:1003403601725>
- Ogata Y (1999) Seismicity analysis through point-process modeling: a review. *Pure Appl Geophys* 155(2–4):471–507. <http://link.springer.com/article/10.1007/s000240050275>
- Ogata Y (2004) Space-time model for regional seismicity and detection of crustal stress changes. *J Geophys Res Solid Earth* 109(B3):B03,308. <http://onlinelibrary.wiley.com/doi/10.1029/2003JB002621/abstract>
- Ogata (2011) Significant improvements of the space-time ETAS model for forecasting of accurate baseline seismicity. *Earth Planets Space* 63(3):217–229. <https://doi.org/10.5047/eps.2010.09.001>, <http://link.springer.com/article/10.5047/eps.2010.09.001>
- Reverso T, Marsan D, Helmstetter A (2015) Detection and characterization of transient forcing episodes affecting earthquake activity in the Aleutian Arc system. *Earth Planet Sci Lett* 412:25–34. <http://www.sciencedirect.com/science/article/pii/S0012821X14007651>
- Rogers G, Dragert H (2003) Episodic tremor and slip on the Cascadia subduction zone: the chatter of silent slip. *Science* 300(5627):1942–1943. <http://science.sciencemag.org/content/300/5627/1942>
- Ruppert D (2002) Selecting the number of knots for penalized splines. *J Comput Graph Stat* 11(4):735–757
- Ruppert D, Wand MP, Carroll RJ (2003) Semiparametric regression. Cambridge University Press, google-Books-ID: Y4uEvXFP2voC
- Schwartz SY, Rokosky JM (2007) Slow slip events and seismic tremor at circum-Pacific subduction zones. *Rev Geophys* 45(3):RG3004. <http://onlinelibrary.wiley.com/doi/10.1029/2006RG000208/abstract>
- Sen MK, Stoffa PL (2013) Global optimization methods in geophysical inversion. Cambridge University Press. <https://books.google.co.in/books?hl=en&lr=&id=FVshAwAAQBAJ&oi=fnd&pg=PR9&dq=Global+Optimization+Methods+in+Geophysical+Inversion&ots=w7e8EJH0lh&sig=CYrivXGlgzjooeSiOpe8LZQN5zA>
- Todd EK, Schwartz SY (2016) Tectonic tremor along the northern Hikurangi Margin, New Zealand, between 2010 and 2015. *J Geophys Res Solid Earth* 121(12):2016JB013,480. <http://onlinelibrary.wiley.com/doi/10.1002/2016JB013480/abstract>
- Utsu T, Ogata Y, Matsu'ura RS (1995) The centenary of the Omori formula for a decay law of aftershock activity. *J Phys Earth* 43(1):1–33. http://www-solid.eps.s.u-tokyo.ac.jp/~hassei/2011/papers/utsu_et_al_1995.pdf
- Wallace LM, Eberhart-Phillips D (2013) Newly observed, deep slow slip events at the central Hikurangi margin, New Zealand: implications for down dip variability of slow slip and tremor, and relationship to seismic structure. *Geophys Res Lett* 40(20):2013GL057,682. <http://onlinelibrary.wiley.com/doi/10.1002/2013GL057682/abstract>
- Wallace LM, Beavan J, Bannister S, Williams C (2012) Simultaneous long-term and short-term slow slip events at the Hikurangi subduction margin, New Zealand: implications for processes that control slow slip event occurrence, duration, and migration. *J Geophys Res Solid Earth* 117(B11):B11,402. <http://onlinelibrary.wiley.com/doi/10.1029/2012JB009489/abstract>
- Wallace LM, Bartlow N, Hamling I, Fry B (2014) Quake clamps down on slow slip. *Geophys Res Lett* 41(24):2014GL062,367. <http://onlinelibrary.wiley.com/doi/10.1002/2014GL062367/abstract>
- Wallace LM, Kaneko Y, Hreinsdóttir S, Hamling I, Peng Z, Bartlow N, D'Anastasio E, Fry B (2017) Large-scale dynamic triggering of shallow slow slip enhanced by overlying sedimentary wedge. *Nat Geosci* 10(10):765–770. <http://www.nature.com/doi/10.1038/ngeo3021>
- Wiemer S (2001) A software package to analyze seismicity: ZMAP. *Seismol Res Lett* 72(3):373–382. <http://srl.geoscienceworld.org/content/72/3/373>
- Zhuang J, Touati S (2015) Stochastic simulation of earthquake catalogs. Community Online Resource for Statistical Seismicity Analysis
- Zhuang J, Ogata Y, Vere-Jones D (2002) Stochastic declustering of space-time earthquake occurrences. *J Am Stat Assoc* 97(458):369–380. <http://www.tandfonline.com/doi/abs/10.1198/016214502760046925>
- Zigone D, Rivet D, Radiguet M, Campillo M, Voisin C, Cotte N, Walpersdorf A, Shapiro NM, Cougoulat G, Roux P, Kostoglodov V, Husker A, Payero JS (2012) Triggering of tremors and slow slip event in Guerrero, Mexico, by the 2010 Mw 8.8 Maule, Chile, earthquake. *J Geophys Res Solid Earth* 117(B9):B09,304. <http://onlinelibrary.wiley.com/doi/10.1029/2012JB009160/abstract>