# Online Seismic Event Picking via Sequential Change-Point Detection


Shuang Li, Yang Cao, Christina Leamon, Yao Xie
H. Milton Stewart School of
Industrial and Systems Engineering
Georgia Institute of Technology
{sli370, yangcao, cleamon3, yxie77}@gatech.edu

Lei Shi, WenZhan Song
Department of Computer Science
Georgia State University
lshi1@student.gsu.edu
wsong@gsu.edu



***Abstract*—Seismic event picking plays a key role in seismology studies. The goal of seismic event picking is to detect the onset of a seismic event, which typically causes an increase in the amplitude of the recorded signal. In this paper, we present a sequential change-point detection method for online seismic event detection, based on the generalized maximum likelihood statistic. We assume that the signals prior to the event are i.i.d. Gaussian random variables with zero mean and known variance, and after the event are i.i.d. Gaussian with zero mean and an *increased* unknown variance. We form a generalized likelihood ratio (GLR) based statistic by replacing the unknown variance with its maximum likelihood estimate, which takes a simple form and has a recursive implementation. An event is detected whenever the statistic exceeds a prescribed threshold. We compare the performance of our GLR procedure relative to the commonly used short-term-average/long-term-average (STA/LTA) algorithm, which is the state-of-the-art for seismic event detection, using large-scale seismic dataset and demonstrate the benefits of our GLR statistic. We also present a joint detection method to utilize the capability of seismic sensors to record signals through three independent channels, to achieve much better detection performance. We also present a method to combine GLR procedure with P-wave and S-wave filtering.**



***Index Terms*—Seismic event detection, sequential change-point detection, generalized likelihood ratio**


## I. Introduction

Seismic event detection and onset time estimation play key roles in geophysical explorations. For example, it is commonly used in determining the state of a volcano and predicting volcanic eruptions (e.g, [3] and [4]), and in travel-time tomography [7]. A widely used detection algorithm is the STA/LTA algorithm presented in [4]. It has been applied in real seismic event detection networks [5], [7] (STA/LTA stands for short-term average and long-term average). STA/LTA is based on a simple moving average of the amplitudes of the signals over a short window that captures short-term signal variations and a long window that captures more long-term background signal levels. However, STA/LTA has limited precision estimating the onset time and may not react to weak seismic signals quickly.

In this paper, we develop a generalized likelihood-ratio (GLR) based sequential change-point detection procedure for quick seismic event detection and precise onset time estimation. It is based on an empirical observation that a seismic event can be well modeled via an increase in the variance. We consider a procedure that detects an *increase* in the variance of i.i.d. Gaussian distributed observations. We assume the variance prior to the change is known (since there is typically a large amount of background data and this parameter can be estimated fairly accurately), and the variance after the change is unknown (since this represents the unknown event). The GLR procedure computes a statistic for every time unit and detects an event when the statistic exceeds a pre-determined threshold. Computation of the statistic is quite simple and can be implemented recursively; we do not have keep the complete history of data. After an event is detected, the maximum likelihood estimate for the onset time is reported. The GLR procedure has superior performance relative to the STA/LTA algorithm, in the expected detection delay and the accuracy in estimating the onset time as demonstrated using simulations and real seismic data.

We also extend our method for joint P-wave and S-wave detection. Seismic waves are a superposition of the P-wave and S-wave, which contain different information about the geological structure (Chapter 11 of [1]). P-waves travel faster, and if we directly apply our method on the original sequence, what we detect is the P-wave. Then we apply a filtering method [5] on the original signal to separate P-waves and S-waves, to estimate the onset of the S-wave as well. We demonstrate this method on a fairly large seismic dataset (Parkfield data).

Change-point detection for seismic event picking using likelihood ratios has been considered in the past. One such example is Section 2.2 in [2], but the procedure studied therein is offline (rather than sequential), and the formulation is also different, where the variance prior to the change and after the change are both unknown. For our problem, we may assume the variance prior to the change is known since we have a large amount of background data. Since our problem is one-sided (meaning we are only interested in an increase in the variance), we may also apply a positive thresholding in the detection statistics and further improve performance.

In seismology literature, other methods have also been considered (but less commonly used) including the wavelet based method [9] (which may not have a simple online implementation) and AR-AIC picking using a joint AR modeling of the noise and the seismic signal and the Akaike Information Criterion (AIC) [6]. AR-AIC requires estimating a more sophisticated AR model which has much higher complexity than our GLR procedure. Statistical properties of these algorithms have not been rigorously studied. On the other hand, due to the simple structure of our GLR statistic, we may characterize its statistical properties which is our ongoing work. These two existing pickers also did not utilizes the sequential change-point framework (i.e., they do not search for the unknown change-point location in the detection statistic).

Our work provides a first attempt to bringing advances in statistical sequential change-point detection to seismic data analysis. Compared to prior work, our contributions include: (1) Presenting a statistical change-point detection approach for online seismic event picking, using the generalized likelihood ratio (GLR) statistic. Although a related approach has been suggested in the literature (Section 1.2.3 in [1] and reference therein), however, only the likelihood ratio based CUSUM statistic is considered, which assumes the post-change parameter to be known. It is known in [8] that CUSUM procedure is sensitive to the error of estimating the parameters. Our GLR statistic assumes the post-change parameter is unknown, which is more robust than the CUSUM statistics. (2) To the best of our knowledge, this work is the first paper in studying the performance of the statistical change-point detection approach, compared with the short-term-average/long-term-average (STA/LTA) method, which is the state-of-the-art in the seismology field. (3) Our work is the first effect in performing such study using modern large-scale modern seismic dataset (Parkfield field dataset). (4) To fully utilizes the capability of seismic sensors to record vibrations in three directions (Z, East-West, and North-South directions), we present a new approach *joint three-channel event detection*. When an event occurs, typically all three channels are affected simultaneously (with different amplitudes of change), and these channels are independent of each other. Hence, this provides diversity for detection and much improves the detection performance.

## II. Problem Formulation

Suppose we are given a sequence of observations $y_i$, $i = 1, 2, \ldots$. Under the hypothesis of no change, the observations follow i.i.d normal distributions with zero mean and a known variance $\sigma_0^2$. Probability and expectation in this case are denoted by $\mathbb{P}^\infty$ and $\mathbb{E}^\infty$, respectively. Alternatively, there exists an *unknown* change-point that occurs at time $\kappa$, $0 \leq \kappa < \infty$, such that the variance of $y_i$ is shifted from $\sigma_0^2$ to some *unknown* parameter value $\sigma_1^2$ for all $i > \kappa$. The probability and expectation in this case are denoted by $\mathbb{P}^\kappa$ and $\mathbb{E}^\kappa$, respectively. The above setting can be formulated as the following hypothesis testing problem:

$$\begin{aligned} H_0: \quad & y_i \sim \mathcal{N}(0, \sigma_0^2), i = 1, 2, \ldots, \\ H_1: \quad & y_i \sim \mathcal{N}(0, \sigma_0^2), i = 1, 2, \ldots, \kappa, \\ & y_i \sim \mathcal{N}(0, \sigma_1^2), i = \kappa+1, \kappa+2, \ldots. \end{aligned} \tag{1}$$

Our goal is to establish a stopping rule (called detection procedure) that stops as soon as possible after a change-point occurs and avoids raising false alarms when there is no change.

## III. Generalized-likelihood ratio (GLR) procedure

We derive a generalized likelihood ratio (GLR) based statistic for detecting the change. Since the observations are independent, for an assumed value of the change-point $\kappa = k$, the log-likelihood for observations by time $t > k$ is given by

$$\begin{aligned} \ell(k, t, \sigma_1^2) &= \sum_{i=k+1}^{t} \log \frac{\frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-y_i^2/(2\sigma_1^2)}}{\frac{1}{\sqrt{2\pi\sigma_0^2}} e^{-y_i^2/(2\sigma_0^2)}} \\ &= \frac{1}{2}\left(\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2}\right)\left(\sum_{i=k+1}^{t} y_i^2\right) - \frac{t-k}{2}\log\frac{\sigma_1^2}{\sigma_0^2}. \end{aligned} \tag{2}$$

Since the post-change variance $\sigma_1^2$ is unknown, we may replace it by its maximum likelihood estimator. Given the current number of observations $t$ and putative change-point time $k$, by setting the derivative of the log-likelihood function (2) with respect to $\sigma_1^2$ to be 0, we may solve for the maximum likelihood estimator:

$$\hat{\sigma}_1^2 = \frac{\sum_{i=k+1}^{t} y_i^2}{t-k}.$$

Substituting this back into (2), we obtain the generalized likelihood ratio (GLR) statistic

$$\ell(k, t, \hat{\sigma}_1^2) = \frac{t-k}{2}\left(\frac{\hat{\sigma}_1^2}{\sigma_0^2} - 1 - \log\frac{\hat{\sigma}_1^2}{\sigma_0^2}\right).$$

We raise an alarm whenever the GLR statistic exceeds a prescribed threshold $b > 0$. Define

$$U_{k,t} \triangleq \frac{\sum_{i=k+1}^{t}(y_i/\sigma_0)^2}{t-k} = \frac{\hat{\sigma}_1^2}{\sigma_0^2}, \tag{3}$$

then the GLR detection procedure is given by

$$T_1 = \inf\left\{t : \max_{t-w \leq k < t-w'} \frac{t-k}{2}(U_{k,t} - \log U_{k,t} - 1) > b\right\}.$$

where $w > 0$ is a use chosen window-size and $w'$ is the minimum number of samples required for detection. By searching only over a window of the past $w$ samples, this reduces the memory requirements to implement the stopping rule, and it also sets a minimum level of change that we want to detect. A practical choice of the window length $w$ should be larger than the largest detection delay we expect. The other parameter $w'$ is the minimum number of observations needed for computing the maximum likelihood estimator for parameters. In our case, we just set $w' = 1$.

Since an increase in the variance of observations will appear when a seismic event occurs, we are only interested in the case where the estimated post-change variance $\hat{\sigma}_1$ is larger than $\sigma_0$. This means $U_{k,t}$ in (3) should be greater than 1; otherwise we should truncate it to be 1. For this consideration, define a statistic

$$V_{k,t} \triangleq \max\{U_{k,t}, 1\}.$$

and consider the following variant of $T_1$:

$$T_2 = \inf\left\{t : \max_{t-w \le k < t-w'} \frac{t-k}{2}(V_{k,t} - \log V_{k,t} - 1) > b\right\}.$$

In addition to detecting the change-point and reporting the seismic event, we are also interested in determining the arrival time of the seismic event. In other words, we are interested in estimating $\kappa$. We determine this by a maximum likelihood estimate when the procedure stops. Assume detection procedure $T_1$ stops at time $t_1$. Then the estimated change-point location $k^*$ of the event determined by applying $T_1$ is given by:

$$k^* = \arg \max_{t_1-w \le k < t_1-w'} \frac{t_1-k}{2}\left(U_{k,t_1} - 1 - \log U_{k,t_1}\right). \quad (4)$$

Similarly, if we apply $T_2$ to detect the change, we obtain the estimated change-point location by replacing $U_{k,t}$ with $V_{k,t}$ in (4).

The statistics involved in both $T_1$ and $T_2$ can be computed recursively:

$$U_{k,t+1} = \frac{t-k}{t-k+1}U_{k,t} + \frac{(y_{t+1}/\sigma_0)^2}{t-k+1},\ t-w \le k \le t-w',$$

where $U_{t,t} \triangleq 0$.

## IV. Performance Evaluations

### A. Performance metrics

We use two standard performance metrics (1) the expected value of the stopping time when there is no change, the average run length (ARL); (2) the expected detection delay (EDD), defined in our case to be the expected stopping time in the extreme case where a change occurs immediately at $\kappa = 0$. Specifically, for any detection procedure $T$, we can use $\mathbb{E}^\infty\{T\}$ and $\mathbb{E}^0\{T\}$ to denote ARL and EDD, respectively.

Large ARL means rare false alarm and EDD provides an upper bound on the expected delay after a change-point until the detection procedure stops. Thus, an efficient detection procedure should have a large ARL and a small EDD. The choice of threshold $b$ as a balancing parameter plays a key role in establishing an efficient detection procedure. Typically, we set $b$ large enough so that ARL is a large number (5000 or 10000) since we prefer a robust detection procedure.

In the problem of seismic event detection, in addition to ARL and EDD, another important metric is the precision of estimating the change-point location which we measure using the mean-square errors.

### B. STA/LTA algorithm

Similar with $T_1$ and $T_2$, the STA/LTA algorithm is also a stopping rule. Let $m$ be number of observations per second. Assume we observe $(y_{(i-1)m+1}, \ldots, y_{im})$ in $i$th seconds, then a measure for seismic amplitude in [4] is given by:

$$R_i \triangleq \frac{\sum_{j=(i-1)m+1}^{im}(y_j - \bar{y}_{i-1})}{m},$$

where $\bar{y}_{i-1}$ is defined as the average observation value in $(i-1)$th seconds. Then STA and LTA is updated based on the equation:

$$X_i = \frac{\sum_{j=0}^{W-1} R_{i-j}}{W},$$

where $W$ denotes the STA or LTA time window size. LTA shows the long term background signal level while the STA responds to short term signal variation. The ratio of STA over LTA is constantly monitored. Once the ratio exceeds a certain prescribed threshold, the STA/LTA algorithm stops and record this stopping time as the starting time of the seismic event. In the following, we assume that the STA window is 5 seconds and LTA window is 30 seconds.

To compare our proposed methods with the classical STA/LTA algorithm, we take the maximum in $T_1$ and $T_2$ every time when we have $m$ observations. But it should be noted that our methods can work at a finer temporal grid than STA/LTA algorithm since they can return one result every $1/m$ seconds. In the following numerical examples, we set $m = 40$. In order words, we obtain one observation every 0.025 seconds.

### C. Simulated data

Through simulation, we compare the expected detection delays for $T_1$, $T_2$ and the STA/LTA algorithm when their ARLs are all approximately $10^5$ seconds (this means the ARL is longer than a day, i.e., we do not expect to make one false alarm in one day). We set the window size $w = 2000$. In the simulation, for each Monte Carlo trial, we generate a sequence of observations of which the first 4000 observations (100 seconds) follow the standard normal distribution and then the variance increases from 1 to $\rho$. Then we run each detection procedure on this sequence and record the number $(t - 100)$ as the simulated EDD, where $t$ is the first stopping time after 100 seconds. In addition, for $T_1$ and $T_2$, we record $k^*$ in (4) as the estimated change-point location. For the STA/LTA algorithm, $k^*$ is just the stopping time. Since $k^*$ may not always be larger than 100, we use mean squared error (MSE) as the measure of performance, which is given by:

$$\text{MSE} = \frac{1}{K}\sum_{i=1}^{K}(k_i^* - 100)^2, \tag{5}$$

where $k_i^*$ is the estimated change-point location at $i$th Monte Carlo trial and $K$ is the number of Monte Carlo trials. We set $K = 1000$ in the following.

**TABLE I:** Simulated **EDDs** for STA/LTA, $T_1$ and $T_2$. The number in parentheses is the standard deviation of the EDDs of the Monte Carlo trials. The trigger thresholds for STA/LTA algorithm, $T_1$ and $T_2$ are 0.966, 11.2 and 9.60, respectively. The time unit is second.

| | $\rho = 1.1$ | $\rho = 1.3$ | $\rho = 1.5$ | $\rho = 2$ |
|---|---|---|---|---|
| STA/LTA | 1499 (1286) | 28.19 (24.51) | 7.49 (4.01) | 3.54 (0.92) |
| $T_1$ | 226.7 (191.0) | 14.87 (6.51) | 6.23 (2.71) | 2.47 (0.99) |
| $T_2$ | 142.5 (120.4) | 12.28 (5.80) | 5.59 (2.57) | 2.22 (0.95) |

Table I shows the simulated EDDs for the STA/LTA algorithm detection procedure $T_1$ and $T_2$ with different values of $\rho$. Compared to the STA/LTA algorithm, $T_1$ and $T_2$ have smaller detection delays, especially for the case when only a small increase occurs in the variance of the signal (e.g, $\rho = 1.1$). Moreover, $T_1$ and $T_2$ seem to be more robust than the STA/LTA algorithm since the standard deviation of EDDs obtained from 1000 Monte Carlo trials of $T_1$ and $T_2$ are basically smaller than that of the STA/LTA algorithm. Similarly, we may say that $T_2$ is better than $T_1$ in detecting the increase in the variance.

In addition to comparing the EDDs, we also compare the accuracy of estimating the change-point location. MSEs obtained from 1000 Monte Carlo trials are shown in Table II, from which we can see that $T_1$ and $T_2$ can estimate the change-point location much more accurately than the STA/LTA algorithm. The reason for this is that STA/LTA regards the stopping time as the starting of the event without any computation for $k^*$ like that in (4).

**TABLE II:** Simulated **MSEs** for estimating the change-point location when applying STA/LTA, $T_1$ and $T_2$. The trigger thresholds for STA/LTA algorithm, $T_1$ and $T_2$ are 0.966, 11.2 and 9.60, respectively.

| | $\rho = 1.3$ | $\rho = 1.5$ | $\rho = 2$ |
|---|---|---|---|
| STA/LTA | 1411 | 63.44 | 13.79 |
| $T_1$ | 12.09 | 2.56 | 0.40 |
| $T_2$ | 13.32 | 3.27 | 0.57 |

### D. Study Parkfield dataset

We evaluate our approach on real seismic data: the continuous GPS monitoring at Parkfield[1] and compare it with the hand-picked event time[2]. We first preprocess the data by filtering out the low (<1Hz) and high (>10Hz) frequency component and then normalize the signals by the estimated signal variance under null. There are a total number of 10 seismic events, and for each event, the signals were recorded by 36 stations, each with three channels. Hence, we have $36 \times 3 \times 10 = 1080$ sequences. Among these sequences, removing some channels that are not affected by the event, we left with 618 sequences for our study. We compare our proposed GLR-based detection procedure $T_1$ and $T_2$ with the commonly used STA/LTA algorithm. Here we use the *window-limited* version of $T_1$ and $T_2$ with window size $w = 2000$. And we choose thresholds in order to make the ARL of all detection procedures to be approximately $10^5$ seconds. See Table III for the comparative results. The detection rate shows the number of sequences that the method uses to successfully detect the occurrence of a seismic event within 2000 observations (50 seconds) after the event happens.

[1]http://earthquake.usgs.gov/monitoring/edm/parkfield/continuous.php
[2]http://www.mit.edu/∼hjzhang/Parkfield/abs.dat

Table III also shows that our proposed detection procedure $T_1$ and $T_2$ also have better performance than the STA/LTA algorithm in practical settings. The EDDs and MSEs for estimating the change-point location obtained by applying $T_1$ and $T_2$ are both smaller than that of the STA/LTA algorithm, and $T_1$ and $T_2$ has larger detection rates. $T_2$ seems to perform better than $T_1$ in estimating the change-point location even if the detection rate is lower.

### E. Joint-detection by combining signals from three channels.

We may also exploit the fact that each seismic sensor records signal in three channels and when event occurs, it is usually observed (to different degrees) in all three channels, to combine signals recorded by three channels. The three channels are Z-channel, East-West channel, and North-South channel, which records the vibrations in three directions, respectively. This *diversity* provides additional power in detection and we found that it can great improve the performance of our detection algorith,.

We try two ways in combining information from three channels for event detection: (1) summing signals from three channels; this boosts SNR and it improves the overall performance including EDD, MSE, and detection rate, as demonstrated in the second panel in Table III, denoted by "combined"; (2) detecting an event whenever any of the three channels detects a change; this has the highest detection rate, but EDD and MSE may not be improved, as demonstrated in the third panel in Table III, denoted by "one-in-three".

## V. P-wave and S-wave separation and detection

The standard sensor equipment of a seismic station can offer the records of seismograms with three channels, the east-west (E), north-south (N), and vertical (Z) channels. From the three channels, P-waves and S-waves can be separated using the technique proposed in [5]. Since P-waves arrive first, we can use the detection procedures proposed above to first determine the arrival time of the

**TABLE III:** EDDs, MSEs for estimating the change-point location and detection rate for STA/LTA, $T_1$ and $T_2$. The number in parentheses is the standard deviation of the EDDs obtained from 618 sequences of observations. The trigger thresholds for STA/LTA algorithm, $T_1$ and $T_2$ are 0.966, 11.2 and 9.60, respectively.

| | EDD (sec) | MSE | Detection Rate |
|---|---|---|---|
| STA/LTA | 6.24 (1.92) | 42.55 | 459/618 |
| $T_1$ | 4.64 (6.43) | 6.29 | 482/618 |
| $T_2$ | 3.36 (2.25) | 4.35 | 455/618 |
| STA/LTA (combined) | 6.04 (0.22) | 36.49 | 492/618 |
| $T_1$ (combined) | 3.77 (5.14) | 3.26 | 504/618 |
| $T_2$ (combined) | 3.29 (3.75) | 2.14 | 489/618 |
| STA/LTA (one-in-three) | 6.31 (2.49) | 45.99 | 540/618 |
| $T_1$ (one-in-three) | 3.81 (5.10) | 6.71 | 558/618 |
| $T_2$ (one-in-three) | 3.04 (2.43) | 5.95 | 537/618 |

P-wave. Then we apply an S-wave filter to separate the P-wave and S-wave. We run again the detection procedure to determine the arrival time of the S-wave.

Next, we introduce how to compute the P-wave filter and S-wave filter at time $t$. Define $N_t \triangleq (y_{t+1}, y_{t+2}, \ldots, y_{t+M})$, where $\{y_i\}_{i=t+1}^{t+M}$ are observations from channel N between time $t+1$ to $t+M$. Similarly, we define $E_t$ and $Z_t$ be the $M$-dimensional vector of observations from channel E and channel Z between time $t+1$ to $t+M$. Then, we compute the covariance matrix at time $t$ as follows:

$$\Sigma_t = \begin{bmatrix} \text{cov}(N_t, N_t) & \text{cov}(N_t, E_t) & \text{cov}(N_t, Z_t) \\ \text{cov}(E_t, N_t) & \text{cov}(E_t, E_t) & \text{cov}(E_t, Z_t) \\ \text{cov}(Z_t, N_t) & \text{cov}(Z_t, E_t) & \text{cov}(Z_t, Z_t) \end{bmatrix}$$

where $\text{cov}(X, Y) = \frac{1}{M}\sum_{i=1}^{M} x_i y_i$ is the covariance between vector $X$ and $Y$. Define the degree of linear polarization $r_t$ as $r_t = 1 - (\lambda_2 + \lambda_3)/(2\lambda_1)$, where $\lambda_1 \geq \lambda_2 \geq \lambda_3$, and $u_1$, $u_2$ and $u_3$ are the eigenvalues and the corresponding eigenvectors of $\Sigma_t$, respectively. Let $u_{11}$ be the first entry of $u_1$. Then the P-wave filter and the S-wave filter at time $t$ are given by:

$$p_t = r_t u_{11}, \quad s_t = r_t(1 - u_{11}),$$

respectively. Then, we multiply $p_t$ to the $t$th observation of Channel N to obtain the filtered P-wave observation. And we multiply $s_t$ to the sum of $t$th observation of Channel E and Channel Z to obtain the filtered S-wave observation. Fig. 1 demonstrates the P-wave and S-wave separation outcome on a real seismic event. In this example, we set $M = 40$.

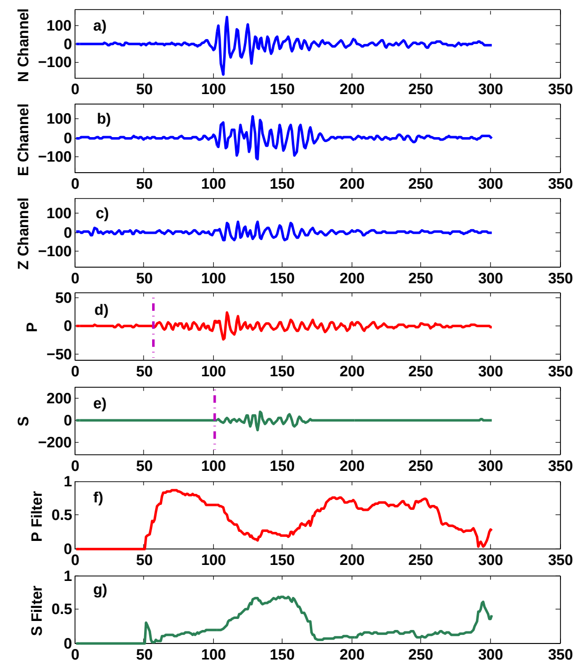


**Fig. 1:** An example of a separated P-wave and S-wave: (a) Channel N, the North component velocity seismogram; (b) Channel E, the East component velocity seismogram; (c) Channel Z, the vertical component velocity seismogram; (d) the separated P-wave by applying the P-wave filter on the signal from Channel Z, where the pink bar marks the estimated change-point time; (e) the separated S-wave by applying the S-wave filter to the sum of signals from Channel E and Channel Z, where the pink bar marks the estimated change-point time; (f) the computed P-wave filter; (g) the computed S-wave filter.

## VI. Conclusion

We formulate the problem of picking seismic events as the detection of increase of variance in Gaussian settings. We present a new generalized likelihood ratio (GLR) based method for online event picking from streaming seismic data, and demonstrate its advantage over the commonly used STA/LTA algorithm for detecting the seismic events. The GLR statistic can be computated recursively, which is convenient for online implementation. We perform extensive numerical study of the performance of the GLR statistic using simulation as we as large-scale seismic dataset (the famous Parkfield dataset). Our comparisons are in terms of two performance metrics: (1) the expected detection delay (EDD) for fixed false-alarm rate, which is represented by average-run-length (ARL) when there is no event, and (2) the mean-square-error (MSE) of estimating the change-point time, which is an important parameter for subsequent seismology study. Moreover, we also present a new joint detection approach, to combine signals from three channels of the seismic sensors to achieve much quicker detection. Finally, we also show how to perform P-wave and S-wave separation jointly with event detection, which is an important tool for seismology study.

## References

[1] Michèle Basseville, Igor V Nikiforov, et al. *Detection of abrupt changes: theory and application*, volume 104. Prentice Hall Englewood Cliffs, 1993.

[2] Jie Chen and Arjun K Gupta. *Parametric statistical change point analysis: With applications to genetics, medicine, and finance*. Springer Science & Business Media, 2011.

[3] RW Decker. State-of-the-art in volcano forecasting. *Bulletin of Volcanology*, 37(3):372–393, 1973.

[4] Elliot T Endo and Tom Murray. Real-time seismic amplitude measurement (rsam): a volcano monitoring and prediction tool. *Bulletin of Volcanology*, 53(7):533–545, 1991.

[5] Zachary E Ross and Yehuda Ben-Zion. Automatic picking of direct p, s seismic phases and fault zone head waves. *Geophysical Journal International*, 199(1):368–381, 2014.

[6] Reinoud Sleeman and Torild van Eck. Robust automatic p-phase picking: an on-line implementation in the analysis of broadband seismogram recordings. *Physics of the Earth and Planetary Interiors*, 111:265–275, 1999.

[7] Wen-Zhan Song, Renjie Huang, Mingsen Xu, Andy Ma, Behrooz Shirazi, and Richard LaHusen. Air-dropped sensor network for real-time high-fidelity volcano monitoring. In *Proceedings of the 7th international conference on Mobile systems, applications, and services*, pages 305–318. ACM, 2009.

[8] Zachary G Stoumbos, Marion R Reynolds Jr, Thomas P Ryan, and William H Woodall. The state of statistical process control as we proceed into the 21st century. *Journal of the American Statistical Association*, 95(451):992–998, 2000.

[9] Haijiang Zhang, Clifford Thurber, and Charlotte Rowe. Automatic P-wave arrival detetion and picking with multiscale wavelet analysis for single-component recordings. *Bulletin of the Seismological Society of America*, 93:1904–1912, 2003.