

## Passive (Micro-) Seismic Event Detection by Identifying Embedded “Event” Anomalies Within Statistically Describable Background Noise

ERICK BAZIW<sup>1</sup> and GERALD VERBEEK<sup>1</sup>

**Abstract**—Among engineers there is considerable interest in the real-time identification of “events” within time series data with a low signal to noise ratio. This is especially true for acoustic emission analysis, which is utilized to assess the integrity and safety of many structures and is also applied in the field of passive seismic monitoring (PSM). Here an array of seismic receivers are used to acquire acoustic signals to monitor locations where seismic activity is expected: underground excavations, deep open pits and quarries, reservoirs into which fluids are injected or from which fluids are produced, permeable subsurface formations, or sites of large underground explosions. The most important element of PSM is event detection: the monitoring of seismic acoustic emissions is a continuous, real-time process which typically runs 24 h a day, 7 days a week, and therefore a PSM system with poor event detection can easily acquire terabytes of useless data as it does not identify crucial acoustic events. This paper outlines a new algorithm developed for this application, the so-called *SEED*<sup>TM</sup> (Signal Enhancement and Event Detection) algorithm. The *SEED*<sup>TM</sup> algorithm uses real-time Bayesian recursive estimation digital filtering techniques for PSM signal enhancement and event detection.

**Key words:** Passive (micro-) seismic monitoring, real-time event detection, Kalman filter, Hidden Markov models.

### 1. Introduction

A passive seismic monitoring (PSM) system is an assembly of hardware and software components designed to acquire and analyse, in real-time, the acoustic signals collected by an array of appropriate seismic transducers. Systems are generally installed in areas where seismicity has been induced by human activity (such as in the vicinity of underground excavations, deep open pits and quarries, around and below large reservoirs where fluids are being injected

into, or removed from, permeable subsurface formations, and adjacent to the sites of large underground explosions (GIBOWICZ and KIJKO, 1994)). Passive seismic monitoring systems are generally classified as microseismic systems if they are designed to measure moment magnitudes less than 0.0  $M_n$ , while systems that record events of stronger magnitude are referred to as macroseismic monitoring systems.

Passive seismic systems are capable of detecting rock failures in the vicinity of underground excavations caused by the sudden release of strain energy resulting from the redistribution of stresses around openings (GIBOWICZ and KIJKO, 1994; TALEBI *et al.*, 1994).

Various hydrocarbon production sites also benefit from seismic monitoring systems during certain phases of production. Primary or secondary extraction or the injection of material into the reservoir to enhance production can cause significant stress changes. These stress changes can result in failures of the overlying strata and the migration of hydrocarbons to aquifers or to the ground surface. Thus, PSM can be used to satisfy environmental concerns, meet regulatory requirements and assess the development of induced fracturing within the reservoir. In addition, passive seismic monitoring systems have been successful in identifying and locating casing failures due to steam stimulation in oil sands (TALEBI *et al.*, 1998).

During filling of hydroelectric or large irrigation reservoirs, changes in regional loading and pore pressures cause significant stress variations within the surrounding rock mass. These can induce a wide range of micro- and macroseismic events, some of which are capable of causing damage to adjacent structures or to the dam itself. PSMs can locate and characterize these potentially hazardous induced events.

<sup>1</sup> Baziw Consulting Engineers Ltd, 3943 West 32nd Avenue, Vancouver, BC V6S 1Z4, Canada. E-mail: info@bcengineers.com

In regions where the level of induced seismicity is high and it is accompanied by significant ambient noise, it is essential that the passive monitoring systems possess the capability of automatically identifying the P- and S-waves generated by seismic events within the noise contaminated raw seismic time series. Reliable automated identification allows for the timely analysis of a large volume of data and the delivery of results to the end user in a useful manner. The ability to locate passive seismic events accurately is directly dependent upon the ability to identify the P- and S-wave responses (phase association) and determine subsequent arrival times (phase picking) (GE and KAISER, 1992).

PSM consists of four elements:

- system specification (e.g., type of sensors, orientation and location) (GE and HARDY, 1988)
- event detection
- source location estimation
- source parameter estimation (e.g., attenuation, seismic moment, source radius, static stress drop, peak particle velocity, seismic energy and failure mechanism) (TALEBI and BOONE, 1998).

The most important of these elements is obviously event detection: the monitoring of seismic acoustic emissions is a continuous, real-time process which typically runs 24 h a day, 7 days a week, and therefore a PSM system with poor event detection can easily acquire terabytes of useless data as it does not identify crucial acoustic events.

PSM event detection consists generally of two steps:

1. Apply a digital filter to the acquired seismic data to increase the signal/noise ratio ( $S/N$ ).
2. Calculate the short term average/long term average ratio (STA/LTA): if this ratio exceeds a user specified threshold then an event is assumed to have occurred and the seismic data is stored.

This paper outlines a PSM real-time event detection filter, which builds upon previous designs (BAZIWI, 2005; BAZIWI and ULRYCH, 2004; BAZIWI and WEIR-JONES, 2002, BAZIWI *et al.*, 2004) in fitting the PSM event detection filter into a Bayesian recursive estimation (BRE) formulation and modeling of the source wavelet as an amplitude modulated sinusoid

(AMS) (BAZIWI, 2005, 2007a, b, 2011; BAZIWI and WEIR-JONES, 2002, 2004; BAZIWI and ULRYCH, 2006; BAZIWI *et al.*, 2004). The event detection filter, the so-called *SEED*<sup>TM</sup> (Signal Enhancement and Event Detection) algorithm, is designed to identify frequency anomalies within a statistically describable noise model in real-time from seismic data acquired from any type of seismic receiver, such as geophones or accelerometers, and utilizes Bayesian recursive estimation to characterize a recorded seismogram in real-time. The real seismic quantities which are characterized are background noise and, if present, a seismic source wavelet “event” (P-wave and/or S-wave). The *SEED*<sup>TM</sup> characterizes these seismic source wavelets as frequency anomalies embedded within background noise, and these anomalies are then compared to a user specified P-wave and S-wave frequency window to assess whether the identified frequency transient is a P-wave, S-wave, or transient noise.

In the original design of the *SEED*<sup>TM</sup> algorithm (BAZIWI and ULRYCH, 2004; BAZIWI and WEIR-JONES, 2002; BAZIWI *et al.*, 2004) a single Kalman filter (KF) formulation was utilized to model the seismogram’s background noise and possible embedded seismic event(s). This KF formulation required a priori knowledge of the dominant frequency of the seismic source wavelet generated when a microseismic “event” occurred. The single KF *SEED*<sup>TM</sup> algorithm would demonstrate poor performance if the source wavelet dominant frequency didn’t match the user specified value. To circumvent this shortcoming, a variant of the *SEED*<sup>TM</sup> algorithm was outlined by BAZIWI (2005). In this filter formulation algorithm a Rao-Blackwellised particle filter (RBPf) and a jump Markov linear Gaussian system (JMLGS) is introduced into the *SEED*<sup>TM</sup> algorithm where changes (i.e., jumps) in the state-space system and measurement equations are due to the occurrences and losses of events within the measurement noise. Significant limitations of the RBPf *SEED*<sup>TM</sup> algorithm are that it introduces considerable complexity into the seismic event detection algorithm, and correspondingly requires a significant increase in computational time. To overcome the above described *SEED*<sup>TM</sup> algorithm limitations, a new formulation is outlined in this paper, with only a single hidden Markov filter applied

on a bank of Kalman filters to obtain robust results within significantly reduced computational time requirements (>70 %).

This paper will discuss this algorithm and provide practical examples to demonstrate that the *SEED*<sup>TM</sup> algorithm provides considerable signal enhancement and event detection advantages when processing PSM seismic data, such as: (1) the ability to identify source wavelet “events” embedded in high variance and correlated noise environments; (2) significant S/N improvement; (3) source wavelet arrival time estimation; (4) the ability to derive noise statistics; and (5) dominant frequency estimation.

## 2. The *SEED*<sup>TM</sup> Algorithm

As stated previously, the *SEED*<sup>TM</sup> algorithm uses real-time BRE digital filtering techniques to analyze the raw PSM data. BRE incorporates the different mathematical tools and concepts of Kalman filter (KF), jump Markov linear Gaussian systems, hidden Markov model (HMM) filter, particle filtering (PF), and Rao-Blackwellised particle filtering (RBPF). Another fundamental component of the *SEED*<sup>TM</sup> is modeling of the source wavelet as an amplitude modulated sinusoid (AMS).

### 2.1. Bayesian Recursive Estimation

In the Bayesian approach to optimal estimation, it is attempted to construct the posterior estimate of the state given all available measurements (ARULAMPALAM *et al.*, 2002). The state or state vector denotes the desired parameters or variables obtained from noisy measurement instrumentation (such as the isolated seismic source wavelet from a noisy seismogram). In general terms, it is desired to obtain estimates of the discretized system equation states  $\mathbf{x}_k$  based on all available measurements up to time  $k$  (denoted as  $\mathbf{z}_{1:k}$ ) by constructing the posterior  $p(\mathbf{x}_k | \mathbf{z}_{1:k})$  and having the initial (prior) pdf of the state  $p(\mathbf{x}_0)$  specified as an initial condition. The posterior pdf allows the conditional mean estimate of the state ( $E[\mathbf{x}_k | \mathbf{z}_{1:k}]$ ) to be calculated. The posterior pdf is represented as posterior = likelihood  $\times$  prior/evidence where the

evidence is the normalizing constant in the denominator.

Bayesian recursive estimation (BRE) refers to the case where a new estimate of the posterior pdf and corresponding conditional mean estimate of the state or state vector is derived for every new measurement. BRE is an optimal filtering technique which is based on state-space, time-domain formulations of physical problems. The terminology, state-space, refers to denoting the desired parameters or variables as states and the physics of the problem (e.g., seismic source wavelet) and corresponding measurement equation(s) (e.g., seismic receiver providing P-wave and S-wave measurements) as the space. The main advantages of utilizing a state-space formulation in describing physical problems is three-fold. (1) Time variance of the system and measurement dynamics and statistics can readily be accounted for. (2) Complicated time variant measurement noise models can easily be incorporated into the measurement equations. (3) The ability to utilize process or system noise to compensate for errors in the mathematical model of the system dynamics.

Application of BRE requires that the dynamics of the system and measurement model which relates the noisy measurements to the system state equations be describable in a mathematical representation and probabilistic form which, with initial conditions, uniquely define the system behaviour.

The potentially nonlinear discrete stochastic equation describing the system dynamics is defined as follows:

$$\mathbf{x}_k = \mathbf{f}_{k-1}(\mathbf{x}_{k-1}, \mathbf{u}_{k-1}) \quad \leftrightarrow \quad p(\mathbf{x}_k | \mathbf{x}_{k-1}) \quad (1)$$

In (1), the vector  $\mathbf{f}_k$  is a function of the state vector  $\mathbf{x}_k$  and the process or system noise  $\mathbf{u}_k$ . It is assumed that (1) describes a Markov process of order one. The system or process noise  $\mathbf{u}_k$  allows the filter designer to incorporate uncertainty within the system model. Notation  $p(\mathbf{x}_k | \mathbf{x}_{k-1})$  denotes the probability of obtaining the state or state vector  $\mathbf{x}$  at time index  $k$  given only the state value at time index  $k - 1$ . This is directly related to the prior within the posterior pdf calculation.

The sampled potentially nonlinear measurement equation is given as:

$$z_k = h_k(x_k, v_k) \leftrightarrow p(z_k|x_k) \quad (2)$$

In (2),  $h_k$  depends upon the index  $k$ , the state  $x_k$ , and the measurement noise  $v_k$  at each sampling time. Notation  $p(z_k|x_k)$  denotes the likelihood of the measurement at time index  $k$  given the state or state vector  $x$  at time index  $k$ . This is directly related to the likelihood within the posterior pdf calculation.

The probabilistic state-space formulation described by (1) and the requirement for updating the state vector estimate based upon the newly available measurements described by (2) are ideally suited for the Bayesian approach to optimal estimation.

BRE is a two step process consisting of prediction and an update model (ARULAMPALAM *et al.*, 2002). In the prediction step the system equation defined by (1) is used to obtain the prior pdf of the state at time  $k$  via the Chapman–Kolmogorov equation, which is given as

$$p(x_k|z_{1:k-1}) = \int p(x_k|x_{k-1})p(x_{k-1}|z_{1:k-1})dx_{k-1} \quad (3)$$

The Chapman–Kolmogorov is derived based upon the transitional densities of a Markov sequence. The update step computes the posterior pdf from the predicted prior pdf and a newly available measurement. The posterior pdf is updated via Bayes’ rule as follows:

$$p(x_k|z_{1:k}) = \frac{p(z_k|x_k)p(x_k|z_{1:k-1})}{p(z_k|z_{1:k-1})} \quad (4)$$

The recurrence equations defined by (3) and (4) form the basis for the optimal Bayesian solution. The BRE of the posterior density can be solved optimally when the state-space equations fit into a Kalman filter formulation or a hidden Markov model. Otherwise, the BRE requires a sub-optimal numerical estimation (e.g., particle filter (ARULAMPALAM *et al.*, 2002; BAZIW, 2007a, b, 2011; DOUCET *et al.*, 2000, 2001)) when deriving the posterior pdf.

### 2.1.1 Kalman Filter

As previously stated, the standard set of KF equations can be implemented as an optimal solution to the BRE when the following conditions are met:  $u_k$  and  $v_k$  are zero mean independent Gaussian white noise

processes,  $f_k$  is a linear function of the state vector and process noise,  $h_k$  is a linear function of the state vector and measurement noise, and the initial estimate of  $x_0$  has a Gaussian distribution (BAZIW, 2007a; ARULAMPALAM *et al.*, 2002; GELB, 1974). The index  $i$  denoted in Table 1 facilitates the implementation of a bank of Kalman filters when implementing a particle filter formulation model (BAZIW, 2007a; ARULAMPALAM *et al.*, 2002; DOUCET *et al.*, 2000, 2001).

In Table 1,  $x_k$  denotes the state to be estimated,  $F_{k-1}$  denotes the state transition matrix which describes the system dynamics,  $u_{k-1}$  denotes the process or system noise (model uncertainty),  $G_{k-1}$  describes the relationship between  $x_k$  and  $u_{k-1}$ , and  $H_k$  defines the relationship between the state and the available measurement (seismogram time series). The implementation of (5) to (14) is outlined in detail by BAZIW (2005, 2007a) and BAZIW and WEIR-JONES (2002, 2004) and GELB, (1974). In general terms, (7) and (9) are used to predict the state and measurement, the innovation (10) is then calculated (difference between the actual measurement and predicted measurement) and the state is updated by adding the predicted value (7) with the weighted innovation (13).

The computational sequence of the KF is outlined as follows:

- A. At time index  $k = 0$ , specify initial conditions  $\hat{x}_0$  (where  $E[x_0] = \hat{x}_0$ ),  $P_0$  (where  $E[(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^T] = P_0$ ) and compute  $F_0$  and  $Q_0$ .

Table 1

KF governing equations for JMLGS

Description	Mathematical representation	Eq.
System equation	$x_k^i = F_{k-1}^i x_{k-1}^i + G_{k-1}^i u_{k-1}^i$	(5)
Measurement equation	$z_k^i = H_k^i x_k^i + v_k^i$	(6)
State estimate extrapolation	$\hat{x}_{k k-1}^i = F_{k-1}^i \hat{x}_{k-1 k-1}^i$	(7)
Error covariance extrapolation	$P_{k k-1}^i = F_{k-1}^i P_{k-1 k-1}^i F_{k-1}^{iT} + G_{k-1}^i Q_{k-1 k-1} G_{k-1}^{iT}$	(8)
Measurement extrapolation	$\hat{z}_k^i = H_{k-1}^i \hat{x}_{k k-1}^i$	(9)
Innovation	$\Delta_k^i = z_k^i - \hat{z}_k^i$	(10)
Variance of innovation	$S_k^i = H_k^i P_{k k-1}^i H_k^{iT} + R_k^i$	(11)
Kalman gain matrix	$K_k^i = P_{k k-1}^i H_k^i (S_k^i)^{-1}$	(12)
State estimate update	$\hat{x}_{k k}^i = \hat{x}_{k k-1}^i + K_k^i \Delta_k^i$	(13)
Error covariance update	$P_{k k}^i = [I - K_k^i H_k^i] P_{k k-1}^i$	(14)

In (5) and (6)  $v_k$  and  $u_k$  are *i.i.d* Gaussian zero mean white noise processes with variances of  $Q_k$  and  $R_k$ , respectively (i.e.,  $v_k \sim N(0, R_k)$  and  $u_k \sim N(0, Q_k)$ )

- B. At time index  $k = 1$ , computer  $\hat{x}_{1|0}$ ,  $P_{1|0}$ ,  $H_1$ ,  $R_1$  and the gain matrix  $K_1$ .
- C. Using the measurement  $z_1$  at time index 1, the best estimates of the state at  $k = 1$  is given by  $\hat{x}_{1|1} = \hat{x}_{1|0} + K_1 \Delta_1$ .
- D. Update the error covariance matrix  $P_{1|1}$ .
- E. At time index  $k = 2$ , a new measurement is  $z_2$  obtained and the computational cycle is repeated.

### 2.1.2 Hidden Markov Model Filter

The Hidden Markov Model Filter (HMM filter (also termed a grid-based filter) has a discrete state-space representation and has a finite number of states. In the HMM filter the posterior pdf is represented by the delta function approximation as follows:

$$p(x_{k-1}|z_{1:k-1}) = \sum_{i=1}^{N_s} w_{k-1|k-1}^i \delta(x_{k-1} - x_{k-1}^i) \quad (15)$$

where  $x_{k-1}^i$  and  $w_{k-1|k-1}^i$ ,  $i = 1, \dots, N_s$ , represent the fixed discrete states and associated conditional probabilities, respectively, at time index  $k - 1$ , and  $N_s$  defines the number of particles utilized. The governing equations for the HMM filter are derived by substituting (15) into the Chapman–Kolmogorov equation (3) and the posterior pdf update Eq. (4). This substitution results in the HMM prediction and update equations which are outlined in Table 2 (ARULAMPALAM *et al.*, 2002). The term particle refers to a possible state value. For example, if the possible dominant frequencies of a sinusoid are 100–300 Hz in increments of 0.5 Hz; then there are (300–100)/

0.5 = 400 possible value or particles. The HMM filter weights each one of these particles as outlined (17) and (18) with the optimal minimum variance estimate of the state or state vector given by summing the weighted particles as outlined in (19).

### 2.1.3 Particle Filter

As stated previously, the recurrence equations defined by (3) and (4) form the basis for the optimal Bayesian solution, and except for the KF and HMM exact solutions the BRE requires a sub-optimal numerical estimation approach. To solve the BRE numerically a new family of filters which rely upon sequential Monte Carlo methods have been made popular within the last decade. This family of new filters are most commonly referred to as particle filters (BAZIWI, 2007a; ARULAMPALAM *et al.*, 2002; DOUCET *et al.*, 2000, 2001).

Similar to the HMM filter, the PF represents the posterior pdf by the delta function approximation, but in this case a randomized grid is utilized for the estimation of the posterior pdf. For the PF, the weights in (15) are obtained using Bayesian importance sampling and a typical PF algorithm is referred to as sequential importance sampling (SIS) (ARULAMPALAM *et al.*, 2002).

## 2.2. Amplitude Modulated Sinusoid

In all variations of the SEED<sup>TM</sup> algorithm, an innovative model of the source wavelet is utilized. This source wavelet model is referred to as the

Table 2

HMM filtering algorithm

Step	Description	Mathematical representation	Eq.
1	Initialization ( $k = 0$ )—initialize particle weights	e.g., $w_k^i \sim 1/N_s$ , $i = 1, \dots, N_s$	(16)
2	Prediction—predict the weights	$w_{k k-1}^i = \sum_{j=1}^{N_s} w_{k-1 k-1}^j p(x_k^i x_{k-1}^j)$	(17)
3	Update—update the weights	$w_{k k}^i = \frac{w_{k k-1}^i p(z_k x_k^i)}{\sum_{j=1}^{N_s} w_{k k-1}^j p(z_k x_k^j)}$	(18)
4	Obtain optimal minimum variance estimate of the state vector and corresponding error covariance	$\hat{x}_k \approx \sum_{i=1}^{N_s} w_{k k}^i x_k^i$ & $P_{\hat{x}_k} \approx \sum_{i=1}^{N_s} w_{k k}^i (x_k^i - \hat{x}_k)(x_k^i - \hat{x}_k)^T$	(19)
5	Let $k = k + 1$ and iterate to step 2		

In the above equations it is required that the likelihood pdf  $p(z_k|x_k^i)$  and the transitional probabilities  $p(x_k^i|x_{k-1}^j)$  be known and specified

amplitude modulated sinusoid (AMS) (BAZIWI, 2005, 2007a, b, 2011; BAZIWI and ULRYCH, 2004, 2006; BAZIWI and WEIR-JONES, 2002; BAZIWI *et al.*, 2004). The AMS is demonstrated to be a highly robust and accurate approximation for many analytical representations of seismic source wavelets such as the exponentially decaying cyclic waveform, the mixed-phase Berlage wavelet, the zero-phase Ricker wavelet, and the zero-phase Klaunder wavelet. In addition, the AMS wavelet has proven very accurate in modeling seismic data acquired during passive seismic monitoring and vertical seismic profiling.

The mathematical representation of the AMS source wavelet is given as

$$x_1(t) = x_2(t) \sin[\omega t + \varphi] \quad (20)$$

where  $x_1(t)$  is an approximation to the seismic source wavelet,  $x_2(t)$  is the seismic wavelet's amplitude modulating term (AMT),  $\omega$  is the dominant frequency of the wavelet, and  $\varphi$  is the corresponding phase.

BAZIWI (2007a, b) outlines the robustness of the AMS model by considering the zero phase Ricker wavelet. The Ricker wavelet is mathematical represented in the time-domain as

$$f(t) = A_0 \left(1 - 2\pi^2 v_M^2 (t - t_0)^2\right) \exp^{-\pi^2 v_M^2 (t - t_0)^2}, \quad (21) \\ t \geq t_0$$

where  $A_0 \equiv$  wavelet maximum amplitude (centered between to flanking lobes),  $v_M \equiv$  dominant or peak frequency of the Ricker wavelet, and  $t_0 \equiv$  wavelet arrival time of maximum amplitude. Although, the

Ricker wavelet has a peak frequency, it doesn't have a specific sinusoidal term and as was shown by BAZIWI (2006, 2007a) the AMS model was able to reconstruct the desired wavelet by applying an appropriate amplitude modulating term.

Another analytical model of the seismic source wavelet is the Berlage wavelet (ALDRIGE, 1990). The Berlage source wavelet is a preferred analytical model of the seismic source wavelet for many researchers due to the fact that mixed-phased source wavelets are readily simulated. For example, Professor Tadeusz Ulrych has utilized Berlage source wavelets extensively in his research and extensive publications.

The Berlage wavelet is defined as

$$w(t) = AH(t)t^n e^{-\alpha t} \cos(2\pi ft + \phi) \quad (22)$$

where  $H(t)$  is the Heaviside unit step function [ $H(t) = 0$  for  $t \leq 0$  and  $H(t) = 1$  for  $t > 0$ ]. The amplitude modulation component is controlled by two factors: the exponential decay term  $\alpha$  and the time exponent  $n$ . These parameters are considered to be nonnegative real constants. Figure 1 illustrates a Berlage wavelet with  $f = 55$  Hz,  $n = 2$ ,  $\alpha = 168$ , and  $\phi = 168^\circ$ . Superimposed upon this Berlage wavelet is a scaled 55 Hz sinusoid with zero crossing at 11.2 ms. As is evident from Fig. 1, the Berlage source wavelet is an amplitude modulated sinusoid

AMINI and HOWIE (2005) utilized a finite difference program (FLAC) to model down-hole seismic source wavelets. BAZIWI (2007a, 2011) illustrates that the AMS source wavelet is an ideal analytical model

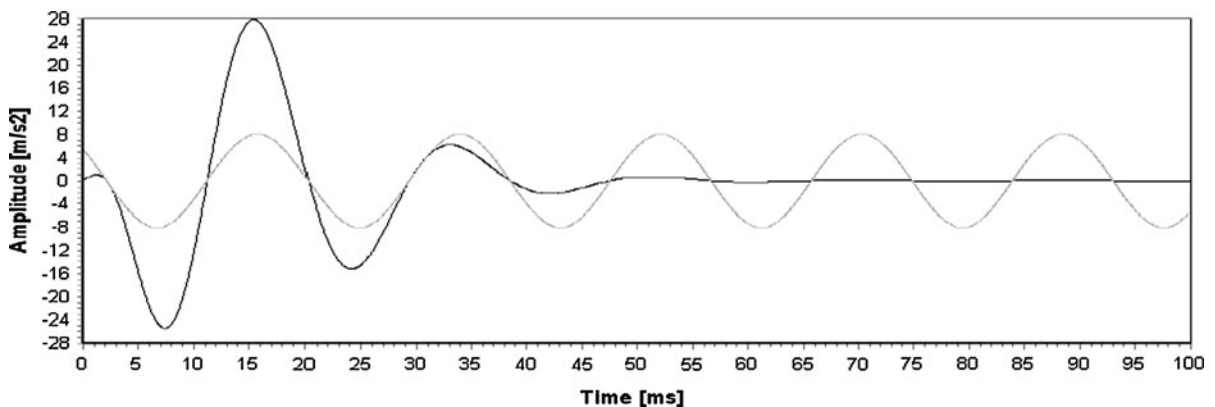


Figure 1  
Berlage wavelet with superimposed 55 Hz sinusoid

for the source wavelet generated by Amini and Howie.

The ability of the analytical AMS source wavelet to model real data is demonstrated with downhole seismic data captured during a seismic cone penetration test (SCPT) (CAMPANELLA *et al.*, 1986, BAZIW, 1993). The seismic sensors utilized were high precision and high bandwidth (1 Hz to 10 kHz) piezoelectric accelerometers, which have highly desirable rise and decay times of approximately 5  $\mu$ s. These fast rise and decay times result in recorded traces where the input of acoustic wavelets and ambient noise are recorded with minimal, if any sensor distortion.

Figure 2 illustrates noisy SCPT data recorded at a depth of 15 m. The high noise energy is due to high frequency rod noise traveling down the steel

extension rods and due to the close radial proximity of the source. Figure 3 illustrates the seismic data shown in Fig. 2 superimposed upon the same seismic trace filtered with a zero phase shift 8th order Butterworth 10–150 Hz bandpass filter applied. Also superimposed upon the filtered seismic trace is a 73 Hz sinusoid. As is evident from Fig. 3, the real SCPT source wavelet can be modeled as amplitude modulated sinusoid.

### 2.3. Gauss–Markov Measurement Noise Model

To facilitate greater source wavelet and measurement noise characterization the measurement noise is modeled as a Gauss–Markov process as opposed to simply being defined by a band of frequencies. By analyzing the autocorrelation and power spectrum of

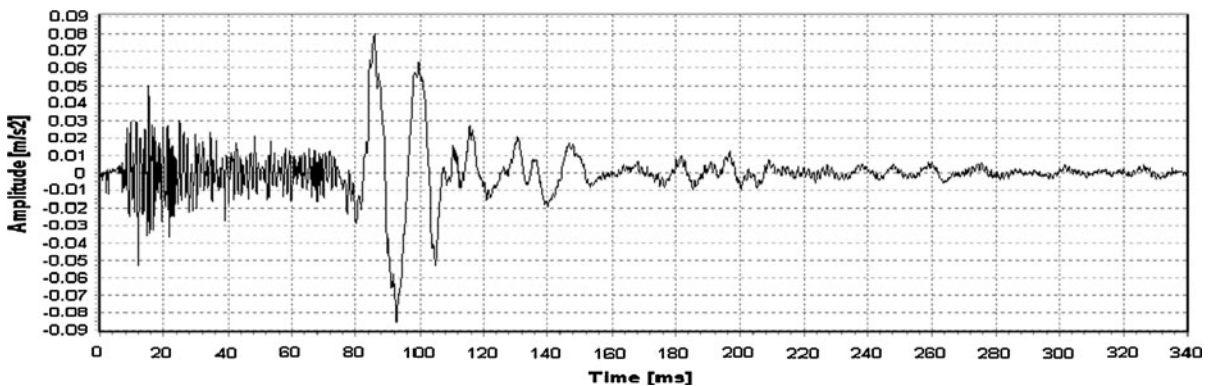


Figure 2  
AMS real data example recorded during a SCPT

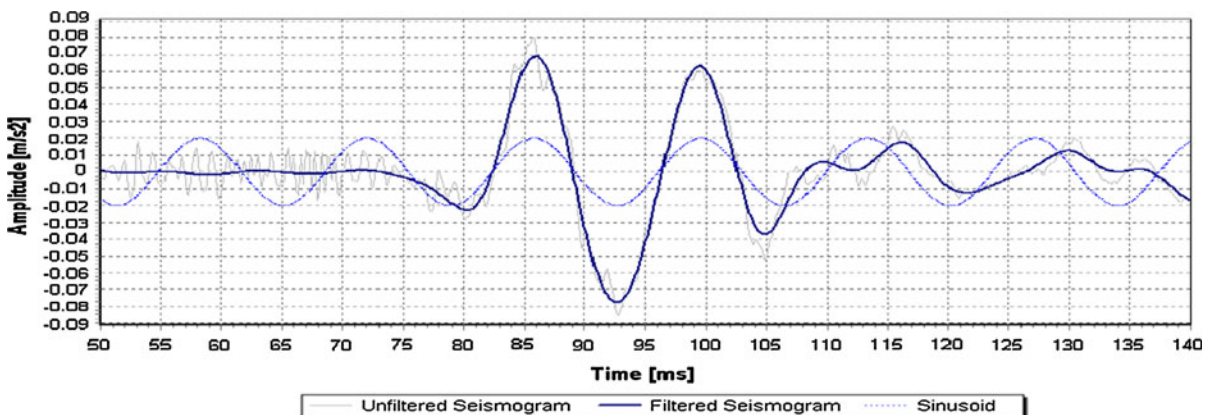


Figure 3  
Seismic trace in Fig. 2 filtered with a 10–150 Hz frequency filtered applied and a 73 Hz sinusoid superimposed

a large number of PSM measurement noise time series it was possible to identify a mathematical model which sufficiently fits the PSM measurement noise process. Figure 4 illustrates the autocorrelation function and power spectral density of common random processes (GELB, 1974) and represents the time correlation of the processes independent of the pdf of the amplitude distribution. A Gauss–Markov process can be used to describe many physical phenomena (BAZIW, 2005; BAZIW and ULRICH, 2004; BAZIW and WEIR-JONES, 2002; BAZIW *et al.*, 2004)) and is a good candidate to model the PSM background noise.

The Gauss–Markov process has a relatively simple mathematical description. As in the case of all stationary Gaussian processes, specification of the process autocorrelation completely defines the process. The variance,  $\sigma^2$ , and time constant,  $T_c$  (i.e.,  $\beta = 1/T_c$ ), define the first-order Gauss–Markov process. The time constant,  $T_c$ , reflects and is directly proportional to the level of correlation between samples. These parameters are derived from the seismic time series by windowing on the noise portion of the trace and calculating the autocorrelation of the ambient noises phenomena (BAZIW, 2005, 2007a, b, 2011; BAZIW and ULRICH, 2006; BAZIW and

WEIR-JONES, 2002; BAZIW *et al.*, 2004). The discrete mathematical equation for a Gauss–Markov process is given as

$$n_{k+1} = a_w n_k + b_w w_k \tag{23}$$

$$a_w = e^{-\beta\Delta} \text{ and } b_w = \sqrt{1 - e^{-2\beta\Delta}}$$

In Eq. (23),  $\Delta$  is the sampling rate and  $w_k$  is a zero-mean, timewise-uncorrelated, unit-variance sequence with a Gaussian probability distribution function.  $n_k$  is therefore a zero-mean, exponentially correlated random variable whose standard deviation is  $\sigma$ . The constant  $a_w$  can have a range of values from  $-1$  to  $+1$ . For a stable variable,  $a_w$  is restricted to values between  $0$  and  $+1$ . For  $a_w \rightarrow 0$ ,  $n_k$  changes rapidly and tends to be uncorrelated from sample to sample. For  $a_w \rightarrow 1$ , the behavior of  $n_k$  becomes more sluggish and it tends to change little from sample to sample.

#### 2.4. SEED<sup>TM</sup> Algorithm Design and Configuration

As previously outlined, the SEED<sup>TM</sup> algorithm uses real-time BRE for PSM signal enhancement and event detection. As a first step the SEED<sup>TM</sup> algorithm applies a bank of finite sinusoids ( $i = 1$  to  $N$ ) with dominant frequencies varying from low ( $f_L$ ) to high ( $f_H$ ) (e.g., 30–430 Hz) and a corresponding frequency resolution  $f_R$ . Each sinusoid represents a possible oscillating term of a source wavelet generated from a seismic event. When a microseismic event occurs the resulting seismic source wavelet is then approximated as an AMS, whereby the sinusoid is modulated by an amplitude modulating term (AMT). As illustrated in Fig. 5, a fixed set of possible sinusoids with corresponding dominant frequencies is specified at the outset. Again, these sinusoids represent the oscillating terms of P-waves and/or S-waves generated from a microseismic event. Next a bank of Kalman Filters are utilized. The KF system equations include the AMT components which are modeled as a two state first order Taylor series with the velocity component represented by a Gauss–Markov process. The KF measurement equations incorporate the sinusoidal components  $\sin(\omega_i t)$  where  $\omega_i = 2\pi f_i$  and  $f_i$  is the dominant frequency. The frequency components are incorporated as states within a HMM filter formulation. The background noise is also included within the

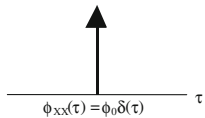
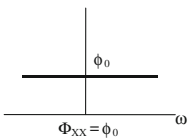
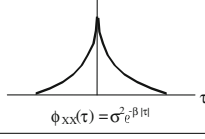
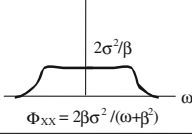
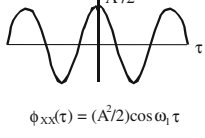
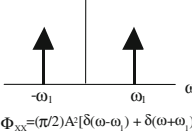
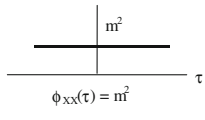
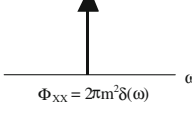
Process	Autocorrelation Function $\phi$	Power Spectral Density $\Phi$
White Noise	 $\phi_{xx}(\tau) = \phi_0 \delta(\tau)$	 $\Phi_{xx} = \phi_0$
Markov Process	 $\phi_{xx}(\tau) = \sigma^2 e^{-\beta \tau }$	 $\Phi_{xx} = 2\beta\sigma^2 / (\omega + \beta^2)$
Sinusoid	 $\phi_{xx}(\tau) = (A^2/2) \cos \omega_1 \tau$	 $\Phi_{xx} = (\pi/2) A^2 [\delta(\omega - \omega_1) + \delta(\omega + \omega_1)]$
Random Bias	 $\phi_{xx}(\tau) = m^2$	 $\Phi_{xx} = 2\pi m^2 \delta(\omega)$

Figure 4 Description of common random processes (13)



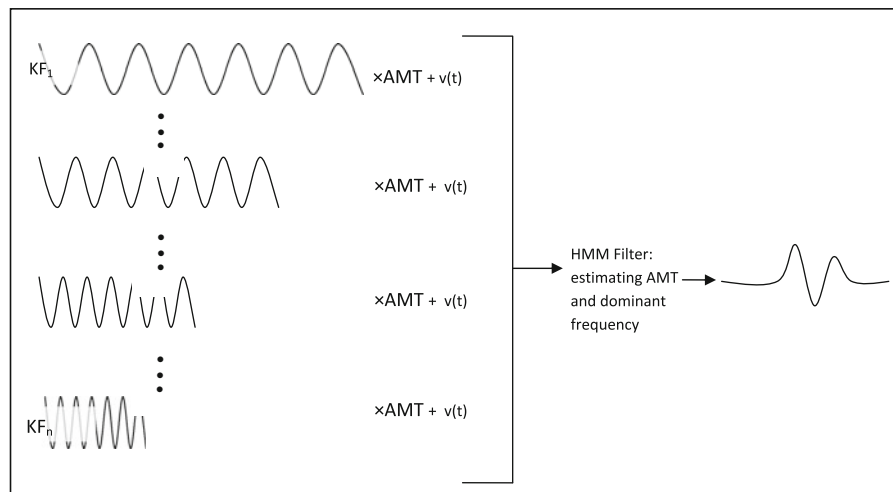


Figure 5  
SEED™ algorithm configuration

KF system equations through the Gauss–Markov process. The required number of KFs and corresponding HMM states is defined as  $N_s = (f_H - f_L)/f_R$ . In addition there is an added KF and HMM state when no source wavelet is present within the seismogram, but only measurement noise (i.e.,  $\omega_0 = f_0 = 0$ ).

For each frequency component,  $f_i$ , in the HMM filter, there is the following three state KF formulation

$$\begin{bmatrix} x_{1k+1}^i \\ x_{2k+1}^i \\ x_{3k+1}^i \end{bmatrix} = \begin{bmatrix} 1 & \Delta & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{bmatrix} \begin{bmatrix} x_{1k}^i \\ x_{2k}^i \\ x_{3k}^i \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & b_2 & 0 \\ 0 & 0 & b_3 \end{bmatrix} \begin{bmatrix} u_{2k+1}^i \\ u_{3k+1}^i \end{bmatrix} \quad (24)$$

In (24) state  $x_1$  denotes the AMT component of the AMS source wavelet, state  $x_2$  denotes the rate of change of state  $x_1$ ,  $x_3$  denotes the Gauss–Markov background noise,  $a_2$ ,  $a_3$ ,  $b_2$  and  $b_3$  are the Gauss–Markov parameters defined in (23),  $u_2$  denotes white Gaussian process noise with mean 0 and variance 1,  $u_3$  denotes white Gaussian measurement noise with mean 0 and variance 1 and  $\Delta$  is the sampling rate.

The SEED™ algorithm models the AMTs (state  $x_1$ ) of the AMS source wavelet with a first order Taylor series approximation. The AMT is forced to be positive by modifying the state estimate extrapolation equation ((7) of Table 1) so that  $\hat{x}_{k|k-1}^i = |F_{k|k-1}^i \hat{x}_{k-1|k-1}^i|$  for state  $x_1$ . The AMT is set positive and the seismogram is also normalized so

that events of varying magnitudes can be processed identically (identical Gauss–Markov parameters for state  $x_2$  for the bank of KFs) and the sinusoidal term of the AMS accounts for source wavelet oscillations. The rate of change term (state  $x_2$ ) is approximated as a Gauss–Markov processes.

As previously outlined, the variance,  $\sigma^2$ , and time constant,  $T_c$ , define the first-order Gauss–Markov process. The time constant terms ( $T_{c2}$  where  $a_2 = e^{-\Delta/T_{c2}}$ ) of state  $x_2$  is an important and very robust parameter within the SEED™ algorithm. Preferably state  $x_2$  results in a smooth trajectory of the amplitude modulation term of the AMS while at the same time allowing for sufficient manoeuvrability so that the AMS source wavelet follows the oscillations of the sinusoid. This parameter is very robust and values of approximately  $a_2 = 0.99$  are found to work very well for modeling a variety of PSM seismic source wavelets.

In specifying the variances of states  $x_2$ , the SEED™ algorithm identifies the approximate maximum rate of change within the normalized seismic time series for a typical seismic source wavelet. The SEED™ algorithm then sets  $\sigma_2^2$  to one-ninth of the square of this maximum. Similar to parameter  $a_2$ , the SEED™ algorithm responds very robustly to minor variation in parameter  $b_2$ .

For each frequency component,  $f_i$ , in the HMM filter, there is the following KF measurement equation

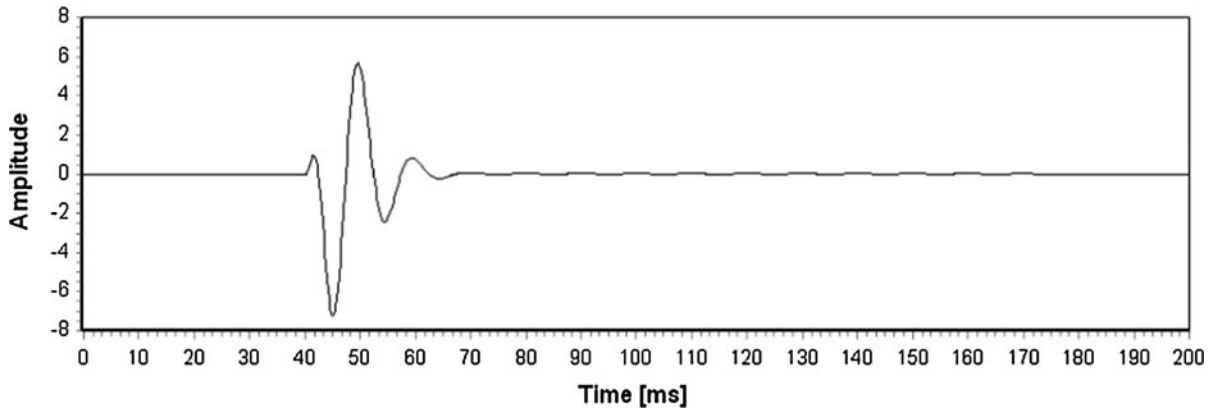


Figure 6  
Berlage source wavelet with dominant frequency of 100 Hz

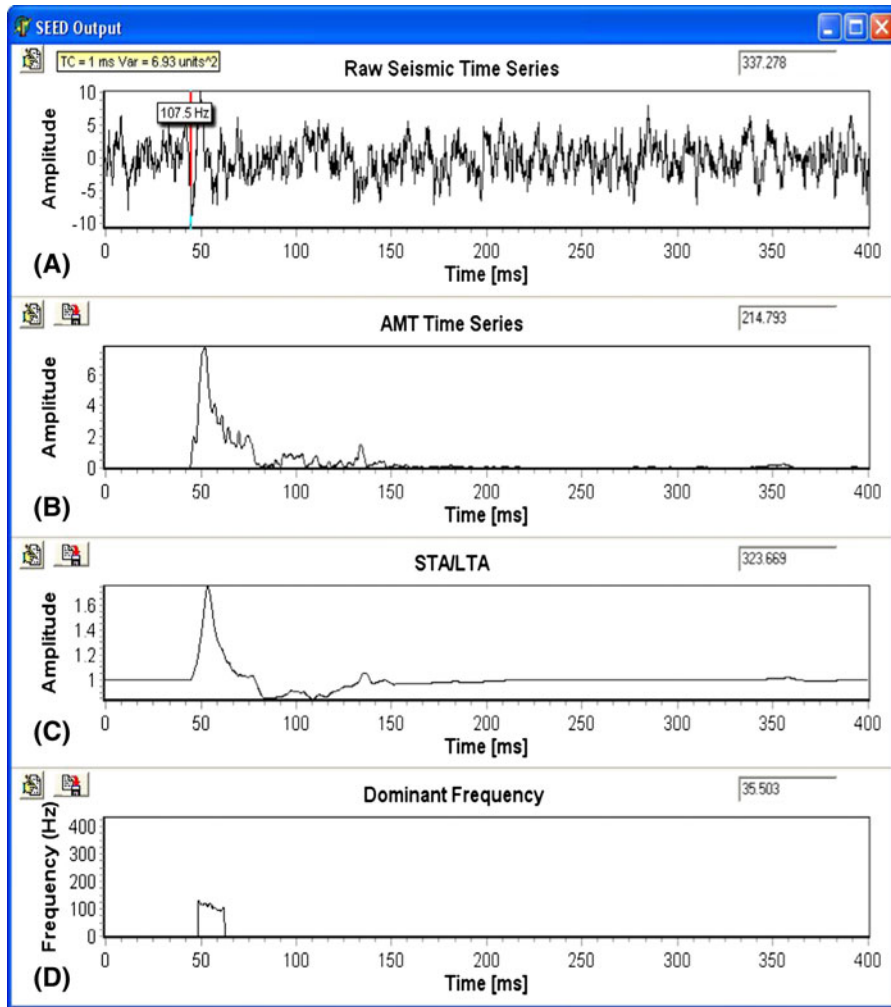


Figure 7

a Synthetic seismogram with source wavelet of Fig. 6 embedded in ambient noise; b derived AMT using the *SEED*<sup>TM</sup> algorithm; c derived STALTA using the *SEED*<sup>TM</sup> algorithm; d estimated frequencies when STALTA threshold of 1.2 exceeded using the *SEED*<sup>TM</sup> algorithm

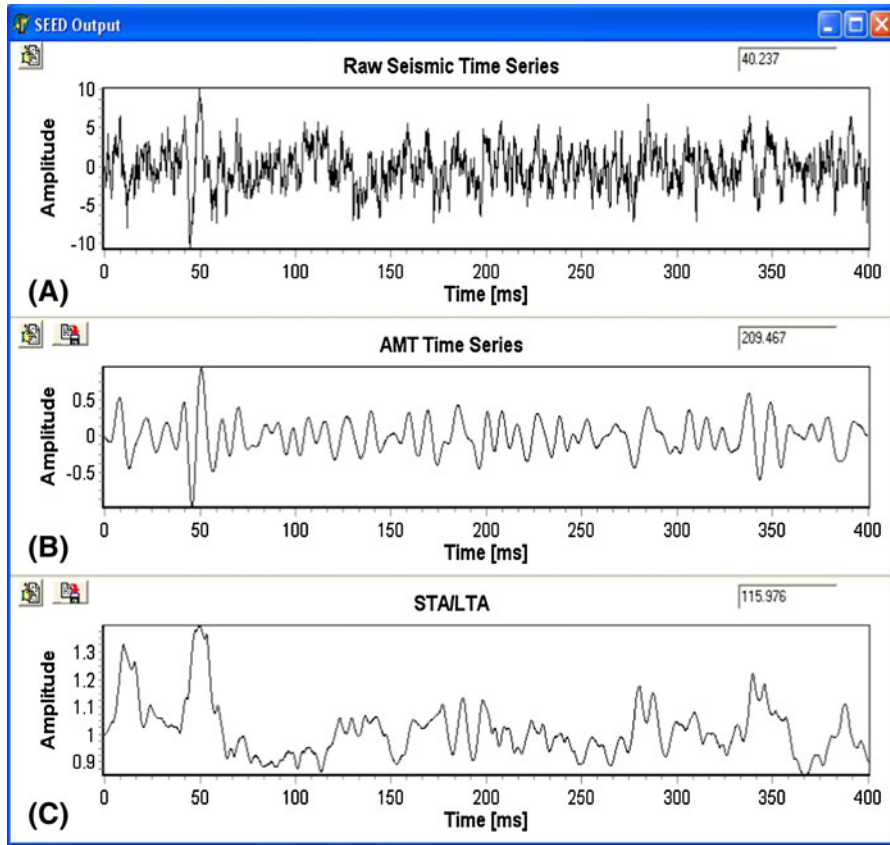


Figure 8

**a** Synthetic seismogram with source wavelet of Fig. 6 embedded in ambient noise; **b** output after applying an eight-order zero phase bandpass (30–150 Hz) to seismogram in (a); **c** derived STALTA of filtered trace

$$z_k^i = x_1^i \sin(\omega_1^i t) + v_k, \quad \text{where } t = k\Delta \quad (25)$$

In (25),  $v_k$  represents the KF measurement noise which is set to a fraction (1%) of the Gauss–Markov noise (state  $x_3$ ) because the Gauss–Markov noise already takes into account background measurement noise (note:  $H_k^i = [\sin(\omega_1^i t) \ 0 \ 0]$  for (9) in Table 1).

The transitional probabilities (i.e.,  $p(x_k^i | x_{k-1}^j)$  or  $p(f_k^i | f_{k-1}^j)$ ) for each state (i.e., frequency component  $f_i$ , and corresponding KF) in the HMM (Table 2) is set relative high for moving from “event” (background noise + source wavelet present) to “non-event” (background noise only present) (i.e.,  $p(f_k^0 | f_{k-1}^j)$ ) and vice versa  $p(f_k^i | f_{k-1}^0)$ . In addition, only the three closest transitional probabilities are specified and calculated, since the dominant frequency is not expected to jump significantly when moving from event to non-event and vice versa. This

assumption also results in a significantly reduced computation time.

The likelihood pdf  $p(z_k | x_k^i)$  in the HMM filter outlined in Table 2 is calculated based upon an assumed Gaussian measurement error as follows:

$$p(z_k | x_k^i) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\left[\frac{(z_k - x_k^i \sin(\omega_1^i t))^2}{2\sigma^2}\right]} \quad (26)$$

where  $\sigma^2 = S_k^i[1, 1]$  ((11) in Table 1 (variance of innovation for state  $x_1$  (AMT)) and  $\hat{x}_{1k}^i$  ((7) in Table 1 (state estimate extrapolation)) are obtained from the associated KF.

The background noise parameters of variance,  $\sigma^2$ , and time constant,  $T_c$ , are automatically derived from the recorded seismic data by windowing on the noise portion and calculating the autocorrelation. For example, on the initiation of the PSM the data acquisition

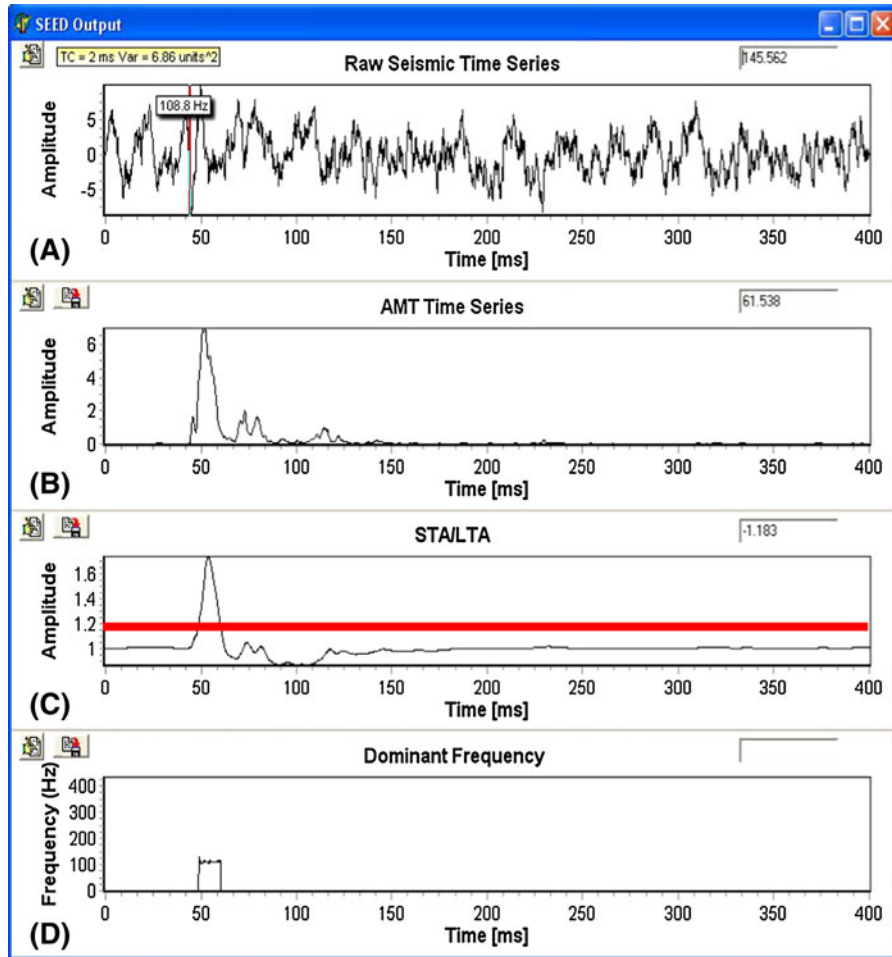


Figure 9

**a** Synthetic seismogram with source wavelet of Fig. 6 embedded in ambient noise; **b** derived AMT using the *SEED*<sup>TM</sup> algorithm; **c** derived STALTA using the *SEED*<sup>TM</sup> algorithm; **d** estimated frequencies when STALTA threshold of 1.2 exceeded using the *SEED*<sup>TM</sup> algorithm

commences acquiring data at the user specified sampling rate  $\Delta$ . A user specified ring buffer of approximately 9,000 points is populated with seismic data and divided into three time windows (e.g., 3,000 points). The variance of each time window is calculated and the window with the lowest variance value is defined to be ambient noise. An auto-correlation is automatically calculated on the ambient noise where the parameters  $\sigma^2$  and  $T_c$  are readily obtained. The investigator should set the data acquisition ring buffer to an appropriate size (time window) so that variability of the background noise is sufficiently characterized.

The ring buffer of data with parameters  $\sigma^2$  and  $T_c$  is then fed into the *SEED*<sup>TM</sup> algorithm for signal processing and event detection. The *SEED*<sup>TM</sup> algorithm calculates the AMT component and the dominant frequency of the AMS source if present. The *SEED*<sup>TM</sup> algorithm must complete filtering and event detection within  $\Delta \times$  ring buffer size (e.g. 9,000 point ring buffer and sampling rate of 5 kHz results in a maximum processing time of 1.8 s) prior to extracting the next set of data from the ring buffer.

The HMM estimated AMT and dominant frequency are calculated from (19) in Table 2 as follows:

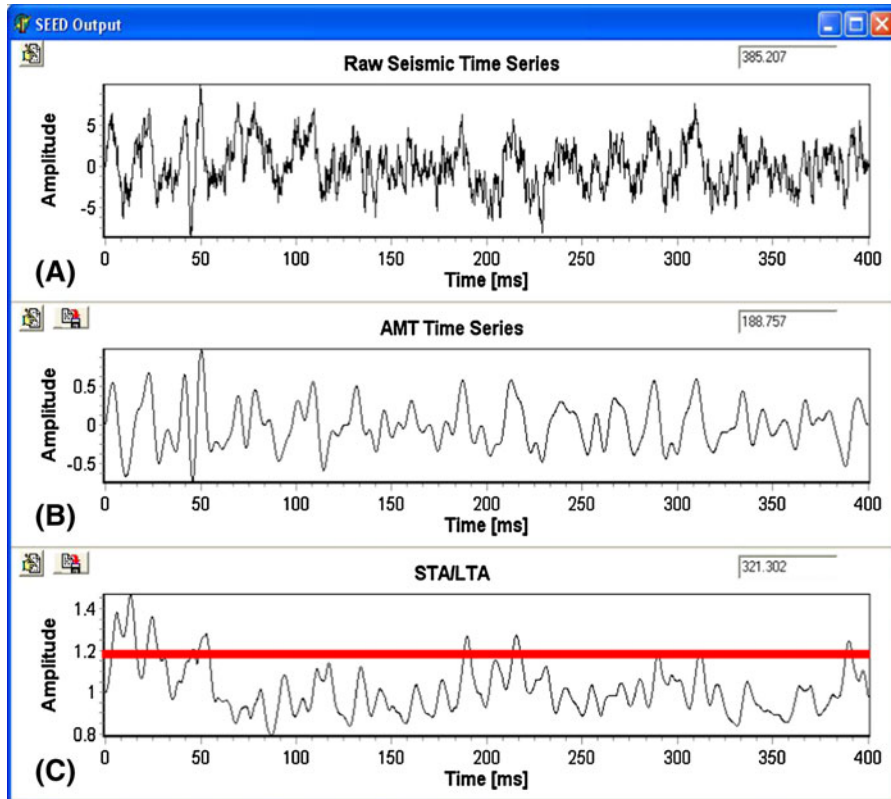


Figure 10

**a** Synthetic seismogram with source wavelet of Fig. 6 embedded in ambient noise; **b** output after applying an eight-order zero phase bandpass (30–150 Hz) to seismogram in (a); **c** derived STA/LTA of filtered trace

$$\hat{x}_{1k} \approx \sum_{i=1}^{N_s} w_{k|k}^i x_{1k}^i \quad (27)$$

$$\hat{f}_k \approx \sum_{i=1}^{N_s} w_{k|k}^i f_k^i \quad (28)$$

In effect, the *SEED*<sup>TM</sup> algorithm acts as a real-time frequency–time–amplitude estimator.

When an “event” is detected the estimated AMT and raw seismic data is stored to file at the user specified sampling rate, sample time and user pre-trigger. There are several parameters which signify an “event” when processing the estimated HMM AMT. The *SEED*<sup>TM</sup> algorithm event detection is based upon the Short Term Average to Long Term Average ratio (STA/LTA) of the estimated AMT exceeding a user specified threshold. The *SEED*<sup>TM</sup> algorithm event detection parameters are outlined as follows:

- **Trigger Channels Required:** the trigger channels required parameter defines how many sensor channels must see an “event” for it to be identified as an overall PSM event.
- **P- or S-wave STA/LTA Parameters.** The *SEED*<sup>TM</sup> algorithm allows for the P-wave and the S-wave trigger to be enabled with differing STA/LTA parameters specified. For each wavelet type the following parameters must be defined:
  - **Threshold Ratio** the threshold for the STA/LTA ratio for the P- or S-wave under analysis.
  - **STA Window [ms]** STA Window length specified in ms for the P- or S-wave under analysis.
  - **LTA Window [ms]** LTA Window length specified in ms for the P-wave under analysis.
- **P- or S-wave Dominant Frequencies** The dominant frequencies input parameters allow the investigator to specify minimum ( $f_{\min}^P$  or  $f_{\min}^S$ )

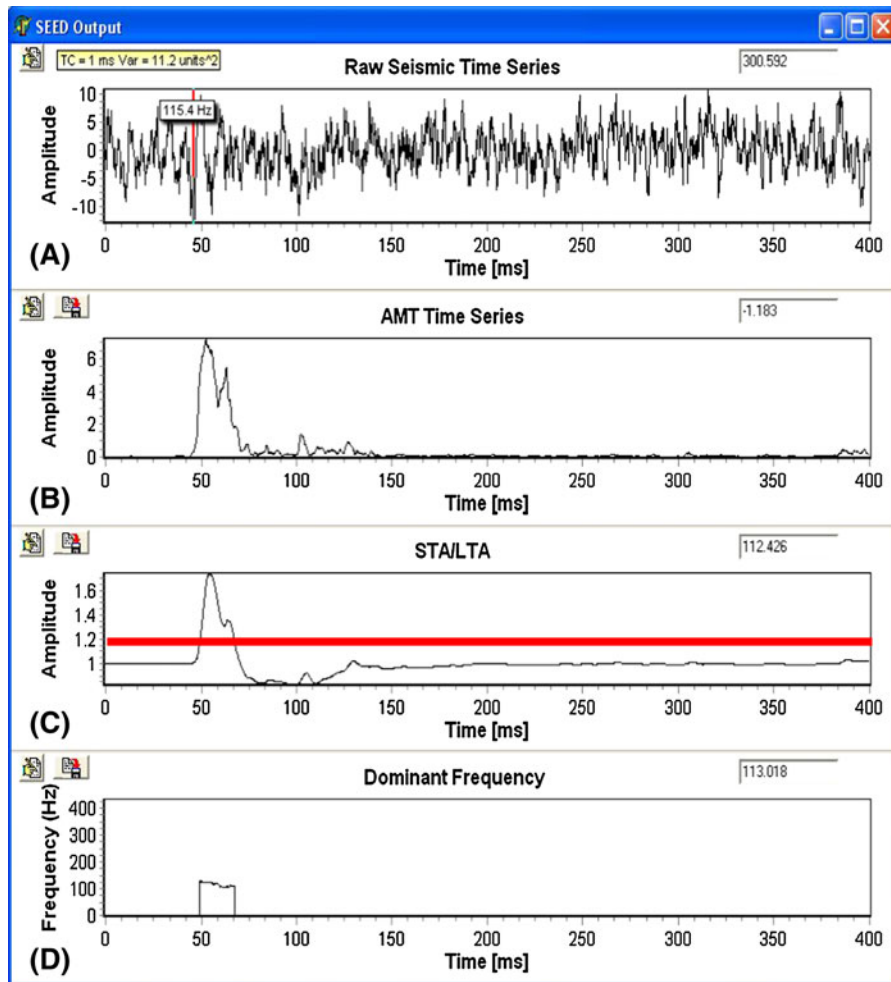


Figure 11

**a** Synthetic seismogram with source wavelet of Fig. 6 embedded in ambient noise with *SEED*<sup>TM</sup> estimated arrival time, dominant frequency and ambient noise statistics illustrated. **b** *SEED*<sup>TM</sup> estimated AMT. **c** *SEED*<sup>TM</sup> estimated STA/LTA. **d** *SEED*<sup>TM</sup> estimated frequencies when STA/LTA threshold of 1.2 exceeded

and maximum ( $f_{\max}^P$  or  $f_{\max}^S$ ) expected values for the P- or S-wave dominant frequency. If the Threshold Ratio is exceeded, the *SEED*<sup>TM</sup> algorithm then verifies that the estimated dominant frequency  $\hat{f}_k$  resides within the P-wave and/or S-wave frequency window.

- *Min Frequency* minimum P- or S-wave dominant frequency ( $f_{\min}^*$ ) in Hz.
- *Max Frequency* maximum P- or S-wave dominant frequency ( $f_{\max}^*$ ) in Hz.
- *Maximum Amplitude Reduction Factor*  
The STA/LTA event detection algorithm is

dependent upon relative amplitudes values (ratio calculation); therefore, there is no indication to what extent the amplitudes have been decreased by applying the *SEED*<sup>TM</sup> algorithm. Consequently it is possible that noise or filter anomalies of very low magnitude (relative to unfiltered seismic data) are present after filtering, due to the fact that there was not a significant source wavelet present. To address this, the Maximum Amplitude Reduction Factor provides an additional level of event detection: if the ratio between the Maximum Amplitude Unfiltered Signal/Maximum AMT exceeds this factor then it is assumed that no “event” has occurred.

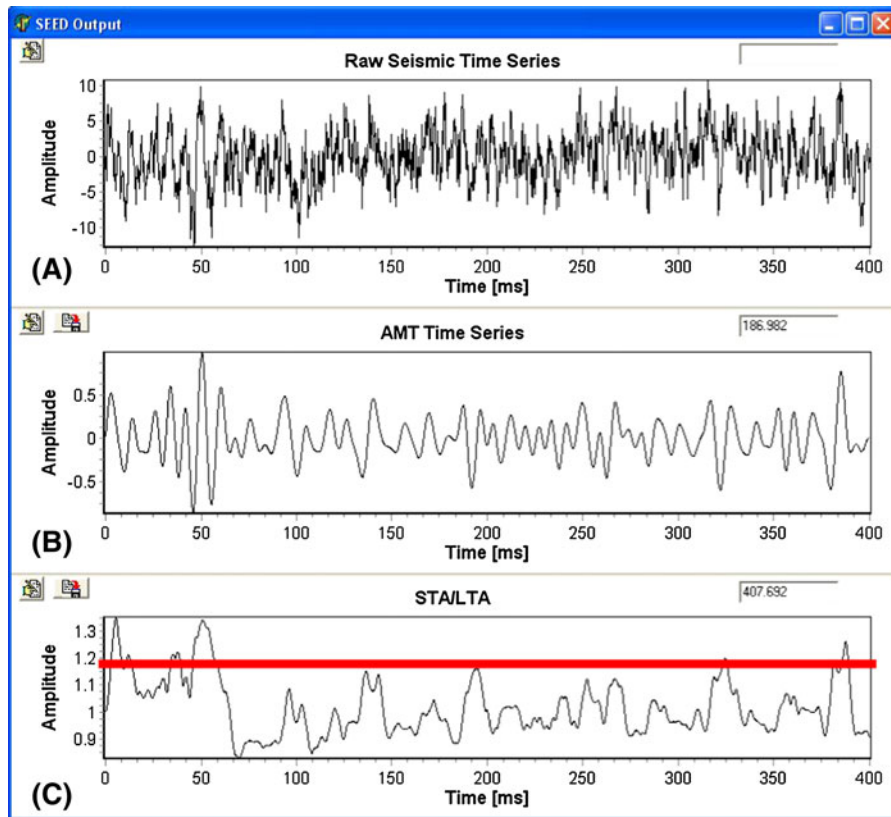


Figure 12

- a** Synthetic seismogram with source wavelet of Fig. 6 embedded in ambient noise with variance of 12 units<sup>2</sup> and time constant of 1 ms.  
**b** Output after applying an eight-order zero phase bandpass (30–150 Hz) to seismogram in (a). STALTA of filtered trace in (b)

This is a very robust parameter and a value between 4 and 8 is typically used.

If the user specified threshold is exceeded, and both the estimated frequency resides within a user specified bandwidth (e.g., P-wave and S-wave bandwidth) and the Maximum Amplitude Reduction Factor is not exceeded, then an “event” is flagged and the time location of the event (time location where STALTA exceeds the user specified Threshold) is defined to be the arrival time of the source wavelet.

### 3. SEED<sup>TM</sup> Performance Results

#### 3.1. Data Simulation

The performance of the SEED<sup>TM</sup> algorithm is assessed by considering challenging synthetic data. Figure 6 illustrates a Berlage source wavelet (BAZIWI, 2005; BAZIWI and WEIR-JONES, 2002, 2004; BAZIWI and

ULRYCH, 2006; BAZIWI *et al.*, 2004) with a dominant frequency of 100 Hz, arrival time of 40 ms and maximum absolute amplitude of 7.2.

#### 3.1.1 Example 1 ( $\sigma^2 = 6 \text{ units}^2$ , $T_c = 1 \text{ ms}$ )

The source wavelet shown in Fig. 6 is embedded within ambient noise with variance of six units<sup>2</sup> and time constant of 1 ms as illustrated in Fig. 7a. The SEED<sup>TM</sup> algorithm is then applied on the noisy seismogram of Fig. 7a where a HMM frequency bandwidth and resolution of 30–430 and 2 Hz is applied, respectively. A STALTA threshold of 1.2 was specified.

The SEED<sup>TM</sup> estimated AMT and STALTA is illustrated in Fig. 7b and c, respectively. Dominant frequencies are estimated when the STALTA ratio exceeds the threshold of 1.2 as shown in Fig. 7d. Figure 7a shows the SEED<sup>TM</sup> estimated noise statistics (variance = 6.9 and time constant = 1 ms), the

averaged dominant frequency estimate (107.5 Hz (true value = 100 Hz)—averaged over time window where STA/LTA ratio exceeds the threshold of 1.2 as is illustrated in Fig. 7d) and arrival time estimate (denoted by vertical bar).

The impressive *SEED*<sup>TM</sup> algorithm results outlined in Fig. 7 are compared with a standard frequency filtering algorithm as illustrated in Fig. 8. Figure 8b shows the seismogram of Figs. 7a and 8a with an eight order digital bandpass (30–150 Hz) filter applied. The STA/LTA of Fig. 8b is shown in Fig. 8c. It is clear from comparing Fig. 8b and c with Fig. 7b and c that the *SEED*<sup>TM</sup> algorithm provided considerable S/N improvement and allowed for dominant frequency and noise statistics estimation.

### 3.1.2 Example 2 ( $\sigma^2 = 9 \text{ units}^2$ , $T_c = 3 \text{ ms}$ )

The source wavelet shown in Fig. 6 is embedded within ambient noise with variance  $\sigma^2$  of 9 units<sup>2</sup> and time constant  $T_c$  of 3 ms as illustrated in Fig. 9a. The *SEED*<sup>TM</sup> algorithm is then applied on this noisy seismogram with a HMM frequency bandwidth and resolution of 30–430 and 2 Hz, respectively. A STA/LTA threshold of 1.2 was specified. The resulting AMT and STA/LTA is illustrated in Fig. 9b and c, respectively. The *SEED*<sup>TM</sup> algorithm also provides the dominant frequencies when the STA/LTA ratio exceeds the threshold of 1.2 as shown in Fig. 9d.

The results using the using the *SEED*<sup>TM</sup> algorithm can be compared with the outcome when using a standard frequency filtering algorithm (applying an eight order digital bandpass (30–150 Hz) filter) as illustrated in Fig. 10. It is clear that *SEED*<sup>TM</sup> algorithm provides a considerable S/N improvement compared to the standard frequency filtering. The other aspect to be considered in this case is the number of times the STA/LTA exceeds the specified threshold: the analysis with the *SEED*<sup>TM</sup> algorithm indicates that there is only 1 event, whereas standard frequency filtering implies that there were possibly 8 events.

### 3.1.3 Example 3 ( $\sigma^2 = 12 \text{ units}^2$ , $T_c = 1 \text{ ms}$ )

The source wavelet shown in Fig. 6 is embedded within ambient noise with variance of 12 units<sup>2</sup> and time

constant of 1 ms as illustrated in Fig. 11a. The *SEED*<sup>TM</sup> algorithm is then applied on the noisy seismogram of Fig. 10a where a HMM frequency bandwidth and resolution of 30–430 and 2 Hz is applied, respectively. A STA/LTA threshold of 1.2 was specified.

The *SEED*<sup>TM</sup> estimated AMT and STA/LTA is illustrated in Fig. 11b and c, respectively. Dominant frequencies are estimated when the STA/LTA ratio exceeds the threshold of 1.2 as shown in Fig. 11d. Figure 11a shows the *SEED*<sup>TM</sup> estimated noise statistics (variance = 11.2 and time constant = 1 ms), the averaged dominant frequency estimate (115 Hz (true value = 100 Hz)—averaged over time window where STA/LTA ratio exceeds the threshold of 1.2 as is illustrated in Fig. 11d) and arrival time estimate (denoted by vertical red bar).

The impressive *SEED*<sup>TM</sup> results outlined in Fig. 11 are compared with a standard frequency filtering algorithm as illustrated in Fig. 12. Figure 12b shows the seismogram of Figs. 11a and 12a with an eighth-order digital bandpass (30–150 Hz) filter applied. The STA/LTA of Fig. 12b is shown in Fig. 12c. It is clear from comparing Fig. 12b and c with Fig. 11b and c that the *SEED*<sup>TM</sup> algorithm provided considerable S/N improvement and allowed for dominant frequency and noise statistics estimation. The other aspect to be considered in this case is the number of times the STA/LTA exceeds the specified threshold: the analysis with the *SEED*<sup>TM</sup> algorithm indicates that there is only one event, whereas standard frequency filtering implies that there were possibly nine events.

## 3.2. Processing Real Data

In this section, the performance of the *SEED*<sup>TM</sup> algorithm is assessed by processing real data acquired with ICP<sup>®</sup> piezoelectric accelerometers manufactured by PCB Piezotronics of New York. The accelerometers have a bandwidth of 1–10 kHz and sensitivity of 100 mV/g. The piezoelectric accelerometers have highly desirable rise and decay times of approximately 5  $\mu$ s. These fast rise and decay times result in recorded traces where the input of acoustic wavelets and ambient noise are recorded with minimal or no sensor distortion.

The seismic data was acquired with a total sampling of 900 ms, pre-trigger of 270 ms, gain of



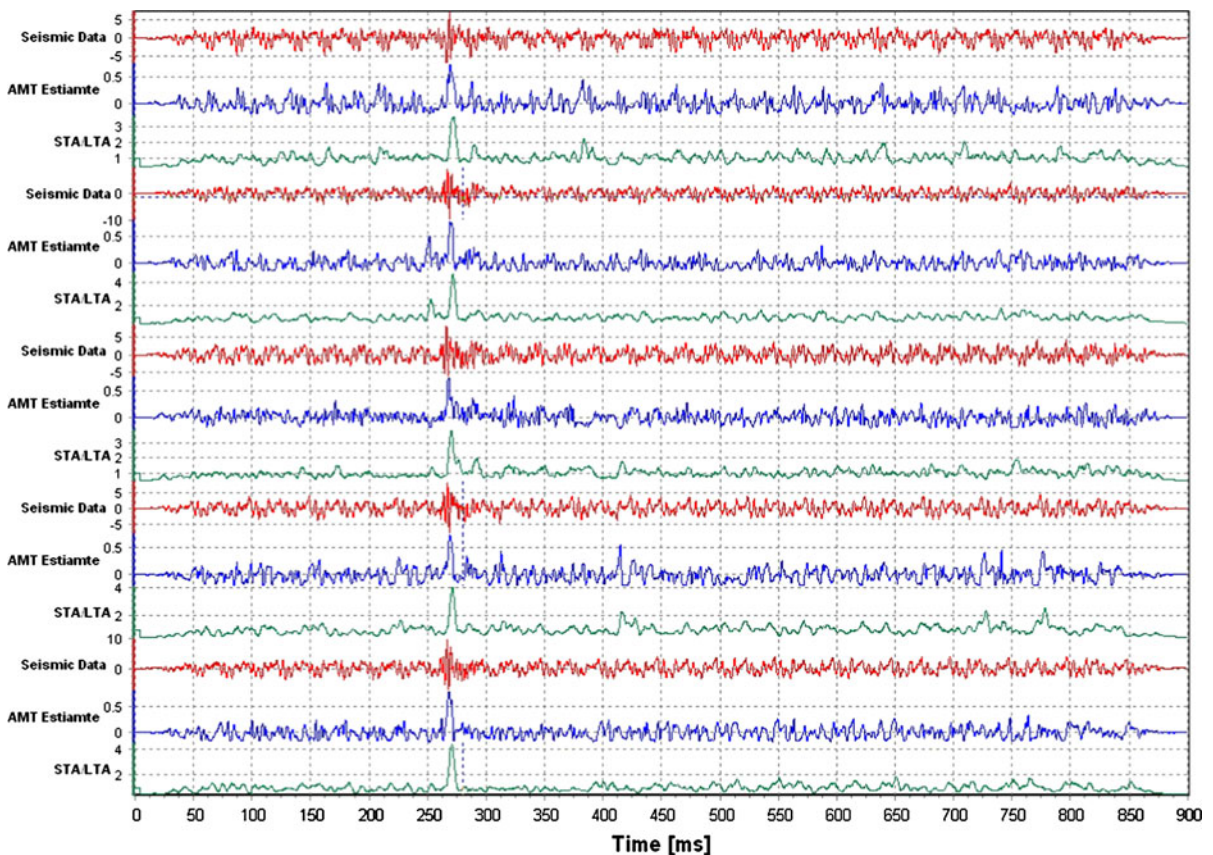


Figure 13

Display of real unfiltered PSM data with pre-trigger at 270 ms, *SEED*<sup>TM</sup> estimated AMT and corresponding STA/LTA time series

6 dB and sampling rate of 5 kHz specified. The *SEED*<sup>TM</sup> algorithm had the following parameters specified:

HMM Filter Parameters:

- Min Frequency—30 Hz
- Max Frequency—430 Hz
- Frequency Resolution—2 Hz.

Trigger Channels Require: 1  
Source Dominant Frequencies

- Min Frequency—160 HZ
- Max Frequency—270 Hz.

Source wavelet STA/LTA Parameters

- Threshold Ratio—2.7.
- STA Window—3.95 ms. Automatically calculated based upon  $0.85 \times$  median source wavelet period ( $STA = 1,000 \times 0.85 / ((270 + 160) \times 0.5)$ ).

- LTA Window (ms)—395 ms. Automatically calculated based upon  $100 \times STA$ .

Maximum Amplitude Reduction Factor: 6

The accelerometers were placed near an industrial fan to simulate a high noise environment. Minor surface impacts (vertical ground impact) were initiated as the PSM system was running so that acoustic event were initiated and the performance of the *SEED*<sup>TM</sup> algorithm assessed when processing real data in a noisy environment.

Figures 13, 14 and 15 illustrate the results from processing real PSM data in a noisy environment. In Fig. 13 the raw seismic traces are shown for five separate events, with the corresponding AMTs and STA/LTA time series. As is shown in Fig. 13, the *SEED*<sup>TM</sup> algorithm was able to extract the “event” in real-time with a high STA/LTA. The estimated *SEED*<sup>TM</sup> algorithm parameters for the five traces

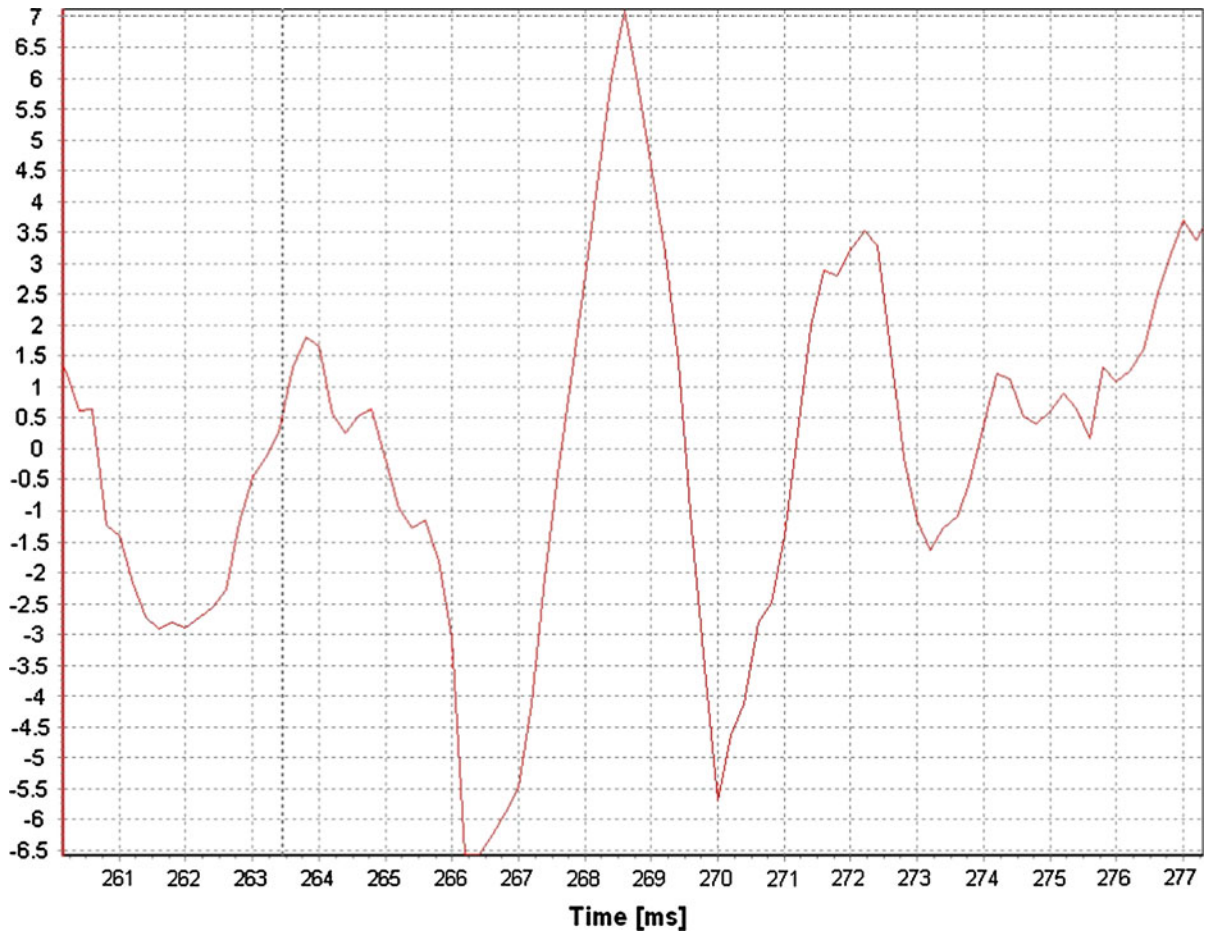


Figure 14

Zoomed-in display of unfiltered seismic time series [trace 1 (top trace)] of Fig. 13. The seismic time series has been zoomed-in around the pre-trigger location (270 ms) so that the peak-to-peak amplitude time offset of approximately 3.8 ms (or dominant frequency of 263 Hz) is shown

(starting from top to bottom) are outlined in Table 3. The background noise in Fig. 13 has an average time correlation between samples of  $T_c = 1.32$  ms or  $a_w = 0.86$  in (23) for a 5 kHz sampling rate. The  $SEED^{TM}$  algorithm provided accurate estimates of the dominant frequencies and Gauss–Markov noise parameters. For example, Fig. 14 shows the top trace of Fig. 13 (i.e., trace 1) zoomed-in around the pre-trigger of 270 ms (“event” time location). The peak-to-peak time offset is estimated at 3.8 ms for a corresponding dominant frequency of 263 Hz. This is very close to the  $SEED^{TM}$  algorithm dominant frequency estimate of 262 Hz. Although the phase and frequency components of the source wavelets and

background noise are similar, the  $SEED^{TM}$  algorithm was able to identify the source wavelets and signify an “event” occurred. Note that the  $SEED^{TM}$  algorithm applies a factor of safety of 20 Hz to the maximum allowed frequency window (e.g., Max Frequency: = Max Frequency + 20 Hz). This is why some of the  $SEED^{TM}$  algorithm dominant frequency estimates exceed 270 Hz.

Figure 15 illustrates the raw PSM seismic time series with an eighth-order zero phase Butterworth bandpass filter applied. The bandpass was set at 160–270 Hz to reflect the desired source wavelet frequency window. As is shown in Fig. 15 very poor STA/LTA time series were obtained with the

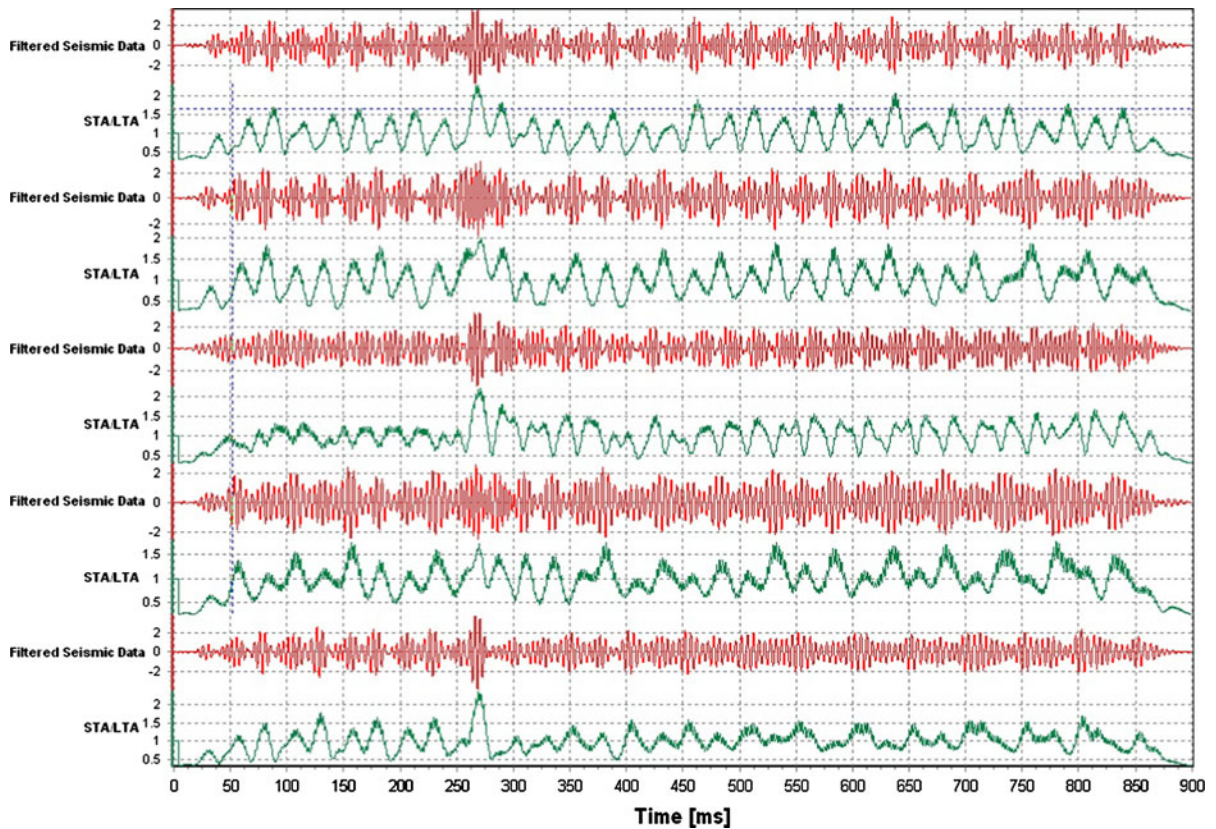


Figure 15

Display of frequency filtered (bandpass of 160–270 Hz) real PSM data and corresponding poor STA/LTA time series

Table 3

*SEED<sup>TM</sup> parameter estimates*

Trace	Dominant frequency (Hz)	Noise $\sigma^2$	Noise $T_c$ (ms)
1	262	2.2	1.2
2	289	2.1	1.4
3	246	2.3	1.2
4	282	2.6	1.4
5	288	2.5	1.4

frequency filtered traces and the user specified threshold was not exceeded for traces 1–5.

#### 4. Conclusions

This paper has outlined a PSM real-time event detection filter which builds upon previous designs in fitting the PSM event detection filter into a Bayesian recursive estimation formulation and modeling of the

source wavelet as an amplitude modulated sinusoid. In general terms, the *SEED<sup>TM</sup>* algorithm quantifies statistically describable background noise into a Gauss–Markov formulation and identifies embedded “event” anomalies in real-time. These anomalies are compared to user specified event detection parameters, such as P-wave and S-wave frequency windows, to assess whether the identified frequency transient is a P-wave, S-wave, or transient noise.

From the discussion and examples provided in this paper, it is obvious that the *SEED<sup>TM</sup>* algorithm provides considerable event detection advantages when processing passive (micro-) seismic data, such as:

- The ability to identify events embedded in high variance and correlated noise environments.
- Significant S/N improvement.
- The ability to derive noise statistics in real time.
- Dominant frequency estimation in real time.

## REFERENCES

- ALDRIGE, D.F. (1990), *The Berlage wavelet*, *Geophysics*, 55(11), 1508–1511.
- AMINI, A. and HOWIE, J.A. (2005), *Numerical Simulation of Downhole Seismic Cone Signals*, *Canadian Geotechnical Journal*, 42(2), 574–586.
- ARULAMPALAM, M.S., MASKELL, S., GORDON, N.J. and CLAPP, T. (2002), *A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking*, *IEEE Transactions on Signal Processing*, 50(2), 174–188.
- BAZIW, E. (1993), *Digital Filtering Techniques for Interpreting Seismic Cone Data*, *Journal of Geotechnical Engineering, ASCE*, 119(6), 98–1018.
- BAZIW, E. (2005), *Real Time Seismic Signal Enhancement Utilizing a Hybrid Rao Blackwellised Particle Filter and Hidden Markov Model Filter*, *IEEE Geosci. Remote Sensing Letters*, 2(4), 418–422.
- BAZIW, E. (2007a), *Application of Bayesian Recursive Estimation for Seismic Signal Processing*. Ph.D. Thesis, Dept. of Earth and Ocean Sciences, University of British Columbia, Vancouver, Canada.
- BAZIW, E. (2007b), *Implementation of the Principle Phase Decomposition Algorithm*, *IEEE Transactions on Geosci. Remote Sensing*, 45(6), 1775–1785.
- BAZIW, E. (2011), *Incorporation of Iterative Forward Modeling into the Principle Phase Decomposition Algorithm for Accurate Source Wave and Reflection Series Estimation*, *IEEE Transactions on Geosci. Remote Sensing*, 49(2), 1775–1785.
- BAZIW, E. and ULRYCH, T.J. (2004), *A Rao-Blackwellised Type Algorithm for Passive Seismic Event Detection*, In *Proceedings of the 8th Annual CDSST Technical Meeting*, Vancouver, Canada, 135–164.
- BAZIW, E. and ULRYCH, T.J. (2006), *Principle Phase Decomposition—A New Concept in Blind Seismic Deconvolution*, *IEEE Transactions on Geosci. Remote Sensing*, 44(8), 2271–2281.
- BAZIW, E. and WEIR-JONES, I. (2002), *Application of Kalman Filtering Techniques for Microseismic Event Detection*, *Pure Appl. Geophys.*, 159, 449–473.
- BAZIW, E., NEDILKO, B. and WEIR-JONES, I. (2004), *Microseismic Event Detection Kalman Filter: Derivation of the Noise Covariance Matrix and Automated First Break Determination for Accurate Source Location Estimation*, *Pure Appl. Geophys.*, 161, 303–329.
- CAMPANELLA, R.G., ROBERTSON, F.T.C. and GILLESPIE, D. (1986), *Seismic Cone Penetration Test*, In: *Proc. IN SITU86*. ASCE, Geot. Spec. Publ., 6, 116–130.
- DOUCET, A., GODSILL, S. and ANDRIEU, C. (2000), *On Sequential Monte Carlo Sampling Methods for Bayesian Filtering*, *Statistics and Computing*, 10(3), 197–208.
- DOUCET, A., de FREITAS, N. and GORDON, N. (2001). *Sequential Monte Carlo Methods in Practice*; (Springer Verlag: New York, NY, USA).
- GE, M. and HARDY, H.R. (1988), *The Mechanism of Array Geometry in the Control of AE/MS Source Location Accuracy*, In *Proc. 29th U.S. Symposium on Rock Mechanics*, Minneapolis, MN, 597–605.
- GE, M. and KAISER, P.K. (1992), *Interpretation of Physical Status of Arrival Picks for Microseismic Source Location*, *Bull. Seism. Soc. Am.*, 80, 1643–1660.
- GELB, A. (1974), *Applied Optimal Estimation* (4th ed., MIT Press: Cambridge, Mass, USA).
- GIBOWICZ, S.J. and KUJO, A. (1994), *An Introduction to Mining Seismology* (Academic Press).
- TALEBI, S. and BOONE, T.J. (1998), *Source Parameters of Injection-Induced Microseismicity*, *Pure Appl. Geophys.* 153, 113–130.
- TALEBI, S., GE, M., ROCHON, P., and MOTTAHED, P. (1994), *Analysis of Induced Seismicity in a Hard-Rock Mine in the Sudbury Basin*, In *Proc. of the 1st North American Rock Mechanics Symposium*, University of Texas at Austin, Texas, June 1–3, 937–944.
- TALEBI, S., NECHTSCHHEIN, S., and BOONE, T.J. (1998), *Seismicity and Casing Failures Due to Steam Stimulation in Oil Sands*, *Pure Appl. Geophys.*, 153, 219–233.

(Received November 15, 2011, revised February 1, 2012, accepted February 27, 2012, Published online April 21, 2012)