Earth Planets Space, **59**, 567–577, 2007

Automatic seismic wave arrival detection and picking with stationary analysis: Application of the KM_2O-Langevin equations

Sho Nakamula[1], Minoru Takeo[1], Yasunori Okabe[2], and Masaya Matsuura[3]

[1]*Volcano Research Center, Earthquake Research Institute, University of Tokyo, Tokyo 113-0032, Japan*
[2]*Department of Mathematical Informatics, Graduate School of Information Science and Technology, The University of Tokyo, Tokyo 113-8656, Japan*
[3]*Department of Mathematics, Faculty of Sciences, Ehime University, Ehime 790-8577, Japan*

(Received June 20, 2006; Revised March 8, 2007; Accepted March 13, 2007; Online published June 27, 2007)

An automatic detection and a precise picking of the arrival times of seismic waves using digital seismograms are important for earthquake early detection systems. Here we suggest a new method for detecting and picking P- and S-wave signals automatically. Compared to methods currently in use, our method requires fewer assumption with properties of the data time series. We divide a record into intervals of equal lengths and check the "local and weak stationarity" of each interval using the theory of the KM_2O-Langevin equations. The intervals are stationary when these include only background noise, but the stationarity breaks abruptly when a seismic signal arrives and the intervals include both the background noise and the P-wave. This break of stationarity makes us possible to detect P-wave arrival. We expand the method for picking of S-waves. We applied our method to earthquake data from Hi-net Japan, and 90% of P-wave auto-picks were found to be within 0.1 s of the corresponding manual picks, and 70% of S-wave picks were within 0.1 s of the manual picks. This means that our method is accurate enough to use as a part of the seismic early detection system.

Key words: Waveform, arrival, picker, KM_2O-Langevin equations, auto-regressive model, P-wave, S-wave.

1. Introduction

We have recently processed a large amount of real time seismic data. Detecting and picking seismic wave arrival quickly is very important for event location and analysis in earthquake early detection systems. In terms of picking seismic signals, manual picking of seismic phases is currently the most accurate method, but it takes much time and unavoidably becomes subjective. Many algorithms have been suggested for the automatic picking of seismic signals. Withers *et al*. (1998) organized previously used methods into four categories: time domain (STA/LTA, Z-statistic), frequency domain (frequency transient), particle motion, and adaptive window length processing. Algorithms based on wavelet analysis (Anant and Dowla, 1997; Zhang *et al*., 2003) and polarization analysis (Vidale, 1986; Reading *et al*., 2001) have also been suggested. The most commonly used algorithm is the autoregressive (AR) model (Yokota *et al*., 1981; Maeda, 1985; Takanami and Kitagawa, 1988; Sleeman and Eck, 1999; Leonard and Kennett, 1999; Leonard, 2000). Methods using the autoregressive model are based on the assumption that seismograms can be divided into two locally stationary intervals at the time of an arrival of seismic signal, with each interval satisfying a different autoregressive process. AR models are fitted to the time series before and after the dividing point which is assumed to be the arrival of seismic signals, and the value of Akaike's information criteria (AIC) is calculated to evaluate the degree of the AR model fitting. By moving this dividing point, the dividing point with the minimum AIC value is judged to be the best point and is the arrival time of seismic signal. However, no check to verify the stationarity of each time series is performed in these methods, and the methods have difficulty in determining the S-wave arrivals if the hypocentral distances are short. Moreover, we must take the time series as it includes just one change of the stationary process prior to the application of the AR-AIC method. Therefore, this method can not use as the "detector" of the phase, though it can use as the "picker". In this paper, we use the word "detector" to denote the method to detect the phase and the word "picker" to denote the method to determine a precise onset time of a detected phase (Allen, 1982). Okabe *et al*. (2003) took the interval of fixed length from seismograms and checked the stationarity of the nonlinear transformed data of this interval. The stationary analysis based on the theory of the KM_2O-Langevin equations (Okabe and Nakano, 1991) was used to check the stationarity of the interval. These authors suggest this algorithm as the method to pick initial phases of earthquakes. The advantage of using this algorithm is that no prior information is required on the data and parametric models, like AR-model, do not have to be defined. We change this algorithm for practical use and present a new method to detect and pick P- and S-wave arrivals automatically by applying the stationary analysis along the time line. We also evaluate the automatic detector and picker of S-wave arrivals when the hypocentral distance is relatively short.

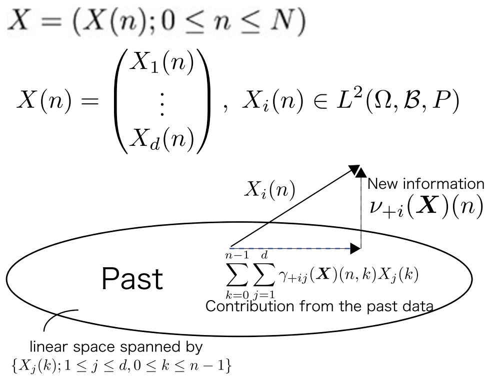


Fig. 1. A general idea of the forward KM_2O-Langevin equations. Data of current time n is constructed by the contribution from the past data plus the information newly given at the time n.

2. Theoretical Background

2.1 Theory for stochastic process

The fundamental principle of our method is the theory of the KM_2O-Langevin equations (Okabe, 1999, 2000; Okabe and Yamane 1998; Matsuura and Okabe, 2001). For given d-dimensional stochastic process, $\mathbf{X} = (X(n); 0 \le n \le N)$, which is square integrable, we can derive the KM_2O-Langevin equations

$$
\begin{aligned}
&X(0) = \nu_+(\mathbf{X})(0) \\
&X(n) = -\sum_{k=0}^{n-1} \gamma_+(n,k)X(k) + \nu_+(n) \quad (1 \le n \le N) \\
&X(N) = \nu_-((X))(0) \qquad\qquad\qquad\qquad\qquad (1) \\
&X(N-n) = -\sum_{k=0}^{n-1} \gamma_-(n,k)X(N-k) + \nu_-(n) \\
&(\nu_\pm(n), \nu_\pm(m)) = \delta_{nm}\mathbf{V}_\pm(n) \quad (0 \le m, n \le N)
\end{aligned}
$$

where $\gamma_\pm(n,k)$ is a $d \times d$ matrix which is determined uniquely from the non-degenerate stochastic process $\mathbf{X}$ and called a forward (resp. backward) KM_2O-Langevin dissipation matrix function. We call $-\sum_{k=0}^{n-1}\gamma_\pm(n,k)X(k)$ (resp. $X(N-k)$) the dissipation term and $\gamma_\pm(n)$ fluctuation term. The matrix function $\mathbf{V}_\pm(\mathbf{X}) = (\mathbf{V}_\pm(n); 0 \le n \le N)$ is obtained as the inner product of $\nu_\pm(n)$: $\mathbf{V}(\mathbf{X})(n) \equiv (\nu_\pm(n), \nu_\pm(n))$. It is to be noted that $(\nu_\pm(n), \nu_\pm(m)) = \delta_{nm}\mathbf{V}_\pm(n) \quad (0 \le m, n \le N)$. We define a system $LM(\mathbf{X})$ and call it the KM_2O-Langevin matrix associated with the non-degenerate stochastic process $\mathbf{X}$. In particular, we put $\delta_\pm(n) \equiv \gamma_\pm(n,0)$. Physically, the dissipation term means the part that can be explained by the previous data, and the fluctuation term means the information newly added at the time n. A general idea of the forward KM_2O-Langevin equations is illustrated in Fig. 1. Same as the forward case, an implication of the backward KM_2O-Langevin equations is that the data of current time $N-n$, $X(N-n)$ is constructed by the contribution from the future data (backward dissipation term) and the information newly given at the time $N-n$ (backward fluctuation term).

We define the local and weakly stationary property for the stochastic process X as follows:

1) The expected value of $\mathbf{X}$ is zero. $(E(X(n)) = 0)$,
2) The covariance matrix function $\mathbf{R}$ of $X(n)$ and $X(m)$ depends only on the relative difference between time n and time m, such that $(X(m), {}^tX(n)) = R(m-n) \quad (0 \le m, n \le N)$.

If $\mathbf{X}$ is a local and weakly stationary process, special relationships exist among the elements of $LM(\mathbf{X})$ (Characterization Theorem for Stationary Property (Okabe, 1999)); that is, $LM(\mathbf{X})$ satisfies Dissipation-Dissipation Theorem (DDT) and Fluctuation-Dissipation Theorem (FDT).

On the other hand, we can uniquely construct the KM_2O-Langevin matrix system $LM(\mathbf{R}) = \{\gamma^R_\pm(n,k), \delta^R_\pm(n), V^R_\pm(m); 0 \le k < n \le N, 0 \le m \le N\}$ that satisfies the (DDT) and (FDT) from any positive definite d × d matrix function $\mathbf{R} = (\mathrm{R}(n); |n| \le N)$ with the Toeplitz condition (Okabe, 2000). This theorem is called the Construction Theorem. Details about the Characterization Theorem, Construction Theorem, (DDT), and (FDT) are explained in Appendix A. Using elements of $LM(\mathbf{R})$, we introduce a pseudo-fluctuation process of $\mathbf{X}$ as follows: $\tilde{\nu}_+(n) = X(n) + \sum_{k=0}^{n-1}\gamma^R_+(n,k)X(k)$. Then, the necessary and sufficient condition for $\mathbf{X}$ to have a weakly stationary property and to have $\mathbf{R}$ as the covariance matrix function is that $\tilde{\nu}(n)$ satisfies the following relation.

$$
\text{[Stationarity Condition]} \quad (\tilde{\nu}(m), {}^t\tilde{\nu}(n)) = \delta_{nm}V^R_+(n), \qquad E(\tilde{\nu}(n)) = 0 \qquad (2)
$$

Based on this, Okabe and Nakano (1991) proposed Test(S) as the method to test the stationarity of a given time series. In Test(S), we use the covariance matrix function of a time series $\boldsymbol{x}$ itself as $\mathbf{R}$ and calculate $LM(\mathbf{R})$. We then check the [Stationarity Condition]. If the [Stationarity Condition] is not satisfied, the time series $\boldsymbol{x}$ is determined to be non-stationary because R used here is the true covariance matrix function of $\boldsymbol{x}$.

Almost all of the methods of time series analysis apply *a priori* parametric statistical models (e.g., ARMA model, AR model) but do not check to verify the preconditions assumed in these models before the analysis (e.g., stationarity for AR model) is performed. In the theory of the KM_2O-Langevin equations, we can obtain the characteristic parameters of the time series, i.e. γ, $\mathbf{V}$ by (DDT), (FDT), (PAC) with no priori assumptions when the stationarity of the data is assured. The stationarity of the time series can be tested by Test(S) before analysis. Therefore, the important characteristics of using this theory are that it requires no priori information about the data and that we do not have to define any parametric models before an execution of the stationary analysis based on the theory. An appropriate model can be extracted directly from the analyzed data in the form of difference equations.

2.2 Application to the real data set: Test(S)

In the above explanation, we developed the theory on the stochastic process. In this section, we extend the theory to real data and describe the framework of Test(S), considering real data as occurrence of the stochastic process $\mathbf{X}$.

A sample covariance matrix function $\mathbf{R} = (R_{jk}(l); 0 \le l \le N,\ 1 \le j, k \le d)$ for a d-dimensional data series $\boldsymbol{x}$ is calculated as

$$R_{jk}(l) \equiv \frac{1}{N+1}\sum_{m=0}^{N-l}(x_j(l+m)-\mu_j)(x_k(m)-\mu_k) \quad (3)$$

where μ_j is the j-th component of the mean vector of $\boldsymbol{x}$. Since a length of real data is finite, calculated sample covariance matrix functions are reliable only for a limited range. If a data length of intervals is $N+1$ ($\boldsymbol{x} = (x(n); 0 \le n \le N)$), we can calculate a reliable $\tilde{LM}$ only for a limited length of $M+1<N+1$; the value of M is obtained empirically to be $M \equiv \max\left(\left[\frac{3\sqrt{N+1}}{d}\right]-1,\ \frac{N+1}{5d}-1\right)$ for d-dimensional data (Okabe and Nakano, 1991; Akaike and Nakagawa, 1988). Because of this reason, we take data pieces whose length is $M+1$ from an interval and test the stationarity of each piece by using $\mathbf{R} = (R(l); 0 \le l \le M)$ calculated from the whole interval. The interval is defined to be stationary when the ratio of data pieces that passed the stationarity check among all the data pieces in the interval exceeds a certain threshold. In this sense, M would act somewhat like the maximum order of the AR-model. However, the apparent resemblance of M to the maximum order is caused simply by the practical circumstance and we do not expect to say the length of contribution from the previous time-series is M. The detailed process of Test(S) is described as follows.

[Step 1] Standardization.

A time series $\boldsymbol{x}$ is standardized. The standardized time series and its sample covariance matrix function are written as $\tilde{\boldsymbol{x}} = (\tilde{x}(n); 0 \le n \le N)$ and $R^{\tilde{x}}$, respectively. Stationarity of $\boldsymbol{x}$ is equivalent to that of $\tilde{\boldsymbol{x}}$. Stationarity of the time series $\tilde{\boldsymbol{x}}$ means that the time series $\tilde{\boldsymbol{x}}$ is a realization of a d-dimensional local and weakly stationary stochastic process $\mathbf{X}$ with $R^{\tilde{x}}$ as its covariance matrix function.

[Step 2] Consideration about the finite length

As mentioned above, we can calculate reliable sample covariance matrix functions only for the limited length of $M+1$. Therefore, we think about data pieces $\tilde{\boldsymbol{x}}^{(s)}(n) \equiv \tilde{x}(s+n)$ $(0 \le n \le M)$ for a fixed number of $s \in \{0, 1, \cdots, N-M\}$.

[Step 3] Calculation of the fluctuation term

From $R^{\tilde{x}}$, we can obtain the sample KM$_2$O-Langevin matrix system $LM(\tilde{x})$ (See Appendix A). Using $LM(\tilde{x})$, the sample forward KM$_2$O-Langevin fluctuation term $\nu_+(\tilde{x}^{(s)})$ is extracted as

$$\nu_+(\tilde{x}^{(s)})(n) \equiv \tilde{x}^{(s)}(n) + \sum_{k=0}^{n-1}\gamma_+(\tilde{x})(n,k)\tilde{x}^{(s)}(k) \quad (0 \le n \le M) \quad (4)$$

[Step 4] Standardization and reformation of the $\nu_+(\tilde{x}^{(s)})$

By taking lower triangular matrices $W_+(n)$, such that $V_+(\tilde{x}^{(s)})(n) = W_+(n)^t W_+(n)$, we standardize the $\nu_+(\tilde{x}^{(s)})$ to $\xi_+^{(s)}(n) \equiv W_+(n)^{-1}\nu_+(\tilde{x}^{(s)})(n)$. The [Stationarity Condition] described in Section 2.1 assurees us that "$\tilde{\boldsymbol{x}}^{(s)}$ is a realization of a local and weakly stationary process with $R^{\tilde{x}}$ as its covariance function" if, and only if, "$\xi_+(\tilde{x}^{(s)})$ realizes a d-dimensional standardized white noise."

[Step 5] reconstruction of 1-D time series $\xi^{(s)}$

All components of $\xi_+(\tilde{x}^{(s)})$ are arranged in one line and the 1-D time series $\xi^{(s)} = (\xi^{(s)}(n); 0 \le n \le d(M+1)-1)$ is constructed as:

$$\xi^{(s)} \equiv (\xi_{+1}^{(s)}(0),\ \xi_{+2}^{(s)}(0),\ \cdots,\ \xi_{+d}^{(s)}(0),\ \xi_{+1}^{(s)}(1),\ \cdots,\ \xi_{+d}^{(s)}(1),\ \cdots,\ \xi_{+1}^{(s)}(M),\ \cdots,\ \xi_{+d}^{(s)}(M)) \quad (5)$$

Using $\xi^{(s)}$, the condition about the $\xi_+(\tilde{x}^{(s)})$ described above is equivalent to following [White noise Condition]. "The time series $\xi^{(s)}$ is a realization of a 1-D standardized white noise stochastic process (say ε_t)." If the [White noise Condition] is satisfied in many case of s, we can conclude the time series $\tilde{\boldsymbol{x}}$ is stationary.

[Step 6] Test the [White noise Condition] of $\xi^{(s)}$

To test the [White noise condition] of $\xi^{(s)}$, we check the normality and orthogonality of $\xi^{(s)}(n)$. The sample mean $\mu^{\xi^{(s)}}$, the sample pseudo-variance $v^{\xi^{(s)}}$, and the sample pseudo-covariance $R^{\xi^{(s)}}(n;m)$ $(0 \le n \le L,\ 0 \le m \le L-n)$ are calculated by following formula

$$\begin{aligned} \mu^{\xi^{(s)}} &\equiv \frac{1}{d(M+1)}\sum_{k=0}^{d(M+1)-1}\xi^{(s)}(k)\,, \\ v^{\xi^{(s)}} &\equiv \frac{1}{d(M+1)}\sum_{k=0}^{d(M+1)-1}\xi^{(s)}(k)^2\,, \\ R^{\xi^{(s)}}(n;m) &\equiv \frac{1}{d(M+1)}\sum_{k=m}^{d(M+1)-1-n}\xi^{(s)}(k)\xi^{(s)}(n+k)\,. \end{aligned} \quad (6)$$

It should be noted that $R^{\xi^{(s)}}(n;0)$ is the sample covariance function of time series $\xi^{(s)}$. A new value $L \equiv [3\sqrt{d(M+1)}]-1$ is introduced for the same reason that we introduced the number M.

The [White noise Condition] of $\xi^{(s)}$ can be written as three criteria.

Mean The sample mean value of $\xi^{(s)}$ distributes around zero, i.e. $\mu^{\xi^{(s)}} \sim 0$;

Variance The sample variance of $\xi^{(s)}$ distributes around 1, i.e. $v^{\xi^{(s)}} - 1 \sim 0$;

Orthogonality The sample covariance of $\xi^{(s)}$ is orthogonal, i.e. $R^{\xi^{(s)}}(n;m) \sim 0$.

We should now formulate these criteria. From the central limit theorem, if the criterion [Mean] is held, $\sqrt{d(M+1)}\mu^{\xi^{(s)}}$ following a normal distribution $N(0,1)$ for sufficiently large M. Therefore, the inequality (C-M) is satisfied approximately at a probability of 0.95, as the 95%

confidence limit of $N(0,1)$.

$$\text{(C-M)} \qquad \sqrt{d(M+1)}\left|\mu^{\xi_i}\right| < 1.96 \tag{7}$$

(C-M) is used as the formulated criterion [Mean] in Test(S). In the same way, the formulated criterion [Variance] in Test(S) can be derived in a form of inequality (C-V) described below. In the case of [Mean], $\sqrt{d(M+1)}\mu^{\xi^{(s)}}$ itself can converge to $N(0,1)$ if the [White noise Condition] of $\xi^{(s)}$ is satisfied. However, in the case of [Variance], sample variance subtracted by 1 and multiplied by $\sqrt{d(M+1)}$, i.e. $\sqrt{d(M+1)}(v^{\xi^{(s)}}-1)$, does not converge to $N(0,1)$ but converges to a Gaussian distribution with average 0 and certain variance σ_i. To evaluate the convergence, we have to calculate the σ_i^2, second moment of $\sqrt{d(M+1)}(v^{\xi^{(s)}}-1)$, so the fourth moment around the average of the white noise process ϵ_i. . But the fourth moment cannot be obtained directly from sample covariance functions. To resolve this, we apply the idea of the t-test, which is often used in the field of statistics (e.g., Snedecor *et al.*, 1989; Student, 1908). A new statistic $(v^{\xi^{(s)}}-1)^{\sim}$ instead of the $v^{\xi^{(s)}}-1$ is introduced as follows:

$$(v^{\xi^{(s)}}-1)^{\sim} \equiv \frac{\sum_{k=0}^{d(M+1)-1}(\xi^{(s)}(k)^2-1)}{\sqrt{\sum_{k=0}^{d(M+1)-1}(\xi^{(s)}(k)^2-1)^2}} = \frac{\frac{1}{\sqrt{d(M+1)}\sigma_i}\sum_{k=0}^{d(M+1)-1}(\xi^{(s)}(k)^2-1)}{\frac{1}{\sqrt{d(M+1)}\sigma_i}\sqrt{\sum_{k=0}^{d(M+1)-1}(\xi^{(s)}(k)^2-1)^2}} \tag{8}$$

Employing the law of large numbers and the central limit theorem, the numerator of Eq. (8) converges to $N(0,1)$ and the denominator of the equation, which includes the square root of unbiased variance of $\xi^{(s)}$, converges to χ distribution of $d(M+1)-1$ degrees of freedom. Therefore, $(v^{\xi^{(s)}}-1)^{\sim}$ will converge to a t distribution of $d(M+1)-1$ degree of freedom. As specified by the t-test, the validness of the condition [Variance] in the 95% confidence limit of $N(0,1)$ for $v^{\xi^{(s)}}-1$ will now be transformed into the following inequality (C-V) under an approximate probability of 0.975 for t-distribution.

$$\text{(C-V)} \qquad \left|(v^{\xi_i}-1)^{\sim}\right| < 2.2414 \tag{9}$$

This inequality (C-V) is used as the criterion [Variance].

Formulation of the criterion [Orthogonality] is more complicated. Similar to the derivation of (C-M) and (C-V), a basic concept of the formulation is to transform $R^{\xi^{(s)}}(n;m)$ into a form that converges to the normal distribution $N(0,1)$. To achieve this, we express $d(M+1)R^{\xi^{(s)}}(n;m)$ as the sum of $R_1^{\xi^{(s)}}(n;m)$ and $R_2^{\xi^{(s)}}(n;m)$, which are the sums of independent stochastic variables. Detailed description about the derivation of the formulated criterion [Orthogonality], an inequality (C-O), and concrete forms of $R_1^{\xi^{(s)}}(n;m)$ and $R_2^{\xi^{(s)}}(n;m)$ are described in Appendix B, and here we show the obtained inequality (C-O) only. (C-O) is satisfied at the probability of 0.95.

$$\text{(C-O)} \quad d(M+1)\left(\sqrt{L_{n,m}^{(1)}}+\sqrt{L_{n,m}^{(2)}}\right)^{-1}\left|R^{\xi_i}(n;m)\right| < 1.96 \tag{10}$$

where $L_{n,m}^{(1)}$ and $L_{n,m}^{(2)}$ are the total of the term number of $R_1^{\xi^{(s)}}(n;m)$ and $R_2^{\xi^{(s)}}(n;m)$, respectively. Concrete forms of $L_{n,m}^{(1)}$ and $L_{n,m}^{(2)}$ are shown in Appendix B. In contrast to the case of [Mean] and [Variance], $R^{\xi^{(s)}}(n;m)$ is in the range of $(n,m;\, 0\le n\le L,\, 0\le m\le L-n)$, and we can test plural times whether $R^{\xi^{(s)}}(n;m)$ satisfies the (C-O) according to the value of n, m. Using this fact, a more precise criterion (C-O)′ is suggested.

(C-O)′ (C-O) is satisfied for more than 90% of $R^{\xi^{(s)}}(n;m)$ $(0\le n\le L,\, 0\le m\le L-n)$ (11)

[Step 7] Test(M), Test(V), Test(O)

If Criteria (C-M), (C-V), and (C-O)′ are satisfied in many cases of s, we can say that the time series $\tilde{\boldsymbol{x}}$ is stationary. We recursively execute [Step 3] to [Step 6] toward $\tilde{\boldsymbol{x}}^{(s)(n)}$ from $s=0$ to $s=N-M$. If at least one of the numbers of data pieces for which (C-M), (C-V), and (C-O)′ do not hold exceeds a certain threshold, the time series is said to be non-stationary. We call these checks of (C-M), (C-V) and (C-O)′ Test(M), Test(V) and Test(O), respectively, and call the ratios of the number of data pieces that do not satisfy the (C-M), (C-V), and (C-O)′ to the total number of data pieces in an interval (i.e. $N-M+1$) the "non-stationarity rate" of Test(M), Test(V), and Test(O), respectively. From numerical simulations toward various synthetic data sets, Okabe and Nakano (1991) suggested that the thresholds of "non-stationarity rate" of Test(M), Test(V), and Test(O) used to define a time series as non-stationary are 0.2, 0.3, and 0.2, respectively. To obtain the thresholds of "non-stationarity rate" of Test(M), Test(V), and Test(O), they carried out repeated experiments toward the large number of data sets whose stationarity or non-stationarity can be determined theoretically, such as uniform random numbers, normal random numbers, tent transformation, the logistic transformation, the transformed data of these data sets by taking first difference, adding noise, and so on. They then obtained the threshold values statistically as the values which can discriminate between stationary data sets and the non-stationary data sets even if the data include the random fluctuations and the error from the finiteness. Details about the experiment are described in Okabe and Nakano (1991). In the case of the seismic records treated here, these values of the thresholds are certainly good criteria of the stationarity. In the bottom panel of Fig. 4, the time series of the non-stationarity rate of Test(V) is plotted. The dotted line in the panel is the threshold value of Test(V), i.e., 0.3. The non-stationarity rate of Test(V) exceeds the threshold when the stationarity of the time series (plotted in the top panel) seems to break.

3. Method

We take an interval of fixed length $(N+1)$ from the observed digital seismogram and apply Test(S). We then move

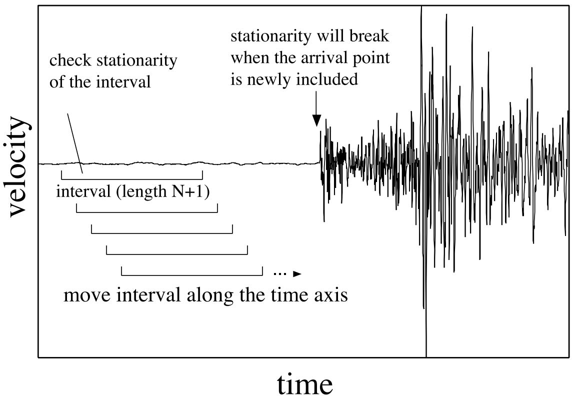


Fig. 2. Concept of our method for determining P-wave arrival by stationary analysis.

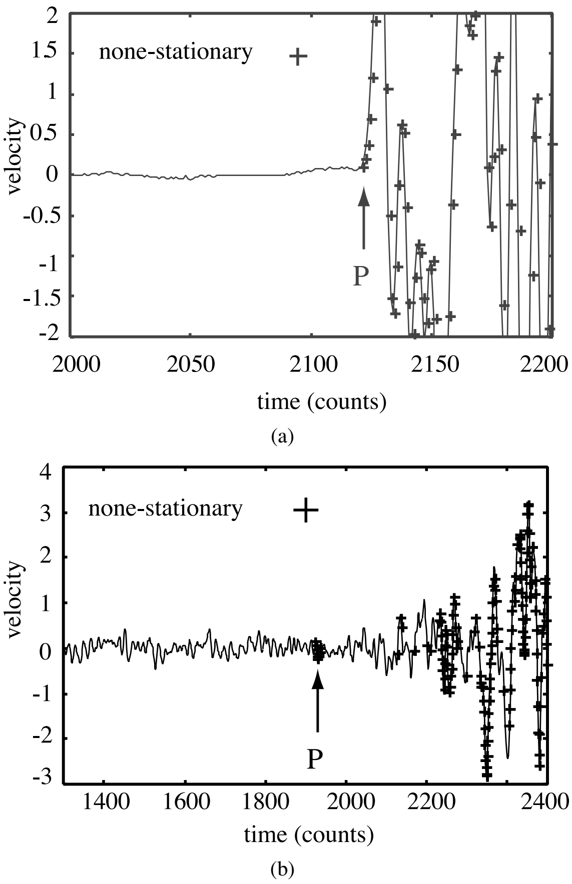


Fig. 3. Typical P-wave arrival as determined by Test(S) with (a) low and (b) high noise. Crosses indicate the last points in the intervals that are determined to be non-stationary.

the interval one data point forward in time (Fig. 2). By moving the interval step by step and applying Test(S) at every step, we can determine how stationarity changes when the last point of the interval is newly included. Stationarity is expected to break abruptly when a point of P-wave arrival is first included.

Figure 3 shows examples of typical P-wave arrivals to which our Test(S) was applied. The sampling rate of the data is 100 Hz. We take the length of the intervals ($N+1$) to be 100, i.e. 1 s. In all of the following figures, the amplitude of the records is normalized to an appropriate size. This normalization does not affect the stationarity of the time series. Figures 3(a) and 3(b) show the results of Test(S) applied to the data with high and low background noise, respectively. If we select $N+1$ to be 100, practical time length $M+1$ would be 20 for 1-D data. In our previous result, the numerical test for using stationary analysis to various basic time-series data is processed and the central limit theorem is sufficiently satisfied if M is about 20. We, therefore, believe that the selected M has enough validity for our problem. In both cases, the interval is stationary while it only includes the background noise, and suddenly becomes non-stationary when the point of P-wave arrival is newly included. The abruptness of the stationarity change depends on the signal/noise ratio and how impulsive the arrival is. When an interval includes the arrival point, the value of the arrival point affects the calculation of all the $\tilde{\nu}, \gamma_{\pm}^{R}$ of pseudo-KM$_2$O-Langevin matrix $LM(\mathbf{R})$through the covariance matrix function $\mathbf{R}$. Consequently, the abrupt break in the stationarity can be seen clearly even if we treat the data with a relatively low signal/noise ratio. We applied Test(S) to a synthetic data set composed of normal random numbers whose mean value is 0 and variance is 1. The data were checked from 0 to 5000 points (=50 s), but no interval was determined to be non-stationary. The length of the interval was 100 points. This result means that the time series are determined to be stationary by Test(S) if the background noise is normal-random. Considering the results, Test(S) can be used for a method to detect and pick the P-wave arrival.

Okabe *et al.* (2003) applied nonlinear transforms of rank 6, i.e. 19 nonlinear transforms, to intervals and made 2-D data sets by combining two of the transformed data. A total of 171 data sets were constructed by this process. They applied Test(S) to the data sets, and if all the data sets were determined to be non-stationary, they called this condition "abnormal" and proposed this criterion as a picker of the initial phase. They called this test Test(ABN). However, a practical algorithm is not constructed in Okabe *et al.* (2003). Taking 19 nonlinear transforms and checking the stationarity of 171 data sets is time-consuming and the application of Test(ABN) as it is to the real time processing system is not practical. We only focus here on the raw time series and construct the detecting and picking algorithm for practical use.

In checking the stationarity of the intervals by Test(S), Test(O) has a special property. Though the non-stationarity rate of Test(O) increases and exceeds the threshold when the interval includes an arrival of P-wave or S-wave, the increase of non-stationarity rate is gradual, and the non-stationarity rate of Test(O) exceeds the threshold not just at the arrival point but a few points after the arrival point. On the other hand, when we pay attention to the last data piece of the interval, i.e. the data set consists of the data from $(N - M + 1)$ to N, the data piece is determined to be non-stationary by Test(O) at a point closer to the arrival point. This benefit is more significant than the increase of the statistical errors introduced by using only one data piece instead of using whole data pieces of the interval.

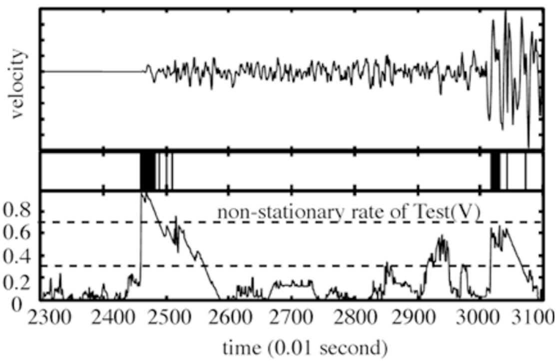


Fig. 4. A seismogram which includes P-wave arrival and the result of Test(V) and Test(O) applied to this seismogram. The top panel shows the waveform with the P-wave arrival of a typical earthquake. The vertical lines in the middle panel mean that the interval was determined to be last-piece-non-stationary by Test(O) when the point above the line was newly included. A curve in the bottom figure shows the non-stationarity rate of the interval determined by Test(V). The value of non-stationarity rate in the interval is plotted at the last point of each interval on this figure. The dotted lines in the bottom panel indicate the minimum and maximum threshold values (0.3 and 0.7) of the non-stationarity rate of Test(V) used in our algorithm

Therefore, we base our picker on this property instead of the non-stationarity rate of Test(O) itself and refer to this property as "last-piece-non-stationary by Test(O)".

Figure 4 shows the non-stationarity rate of Test(V) and the last points of each interval determined to be last-piece-non-stationary by Test(O). At the P-wave arrival, the non-stationarity rate of Test(V) increases sharply, and the interval is determined to be last-piece-non-stationary by Test(O) when the interval includes the arrival point. The picked point is closer to the real arrival point when we use the non-stationary rate of Test(V) for the determination, while less misjudgments for picking occur for the result derived from the last-piece-non-stationarity by Test(O). The non-stationarity rate of Test(V) also increases at points other than the point of real arrival. For this comparison, we constructed our picker so that we detect the arrival point roughly by Test(O) and searched for the accurate arrival point by Test(V) in the vicinity of the roughly estimated point. Here we searched the accurate arrival in a range of 10 points to the roughly estimated point. If the non-stationarity rate of Test(V) exceeds the threshold (we used 0.7 as the threshold here) at a certain point in this time range, we go back along the time line from this point and search the interval which can be regarded as stationary. The last point of this interval is defined as the precise arrival. The non-stationarity rate of Test(M) has always been small and does not require further attention. An outline of our method is shown in Fig. 5.

Using this method, we can detect the phase and determine the precise arrival of the phase (picking) at the same time. This makes real time source location much easier. In this paper, we mainly focus on the part of the picker of this method.

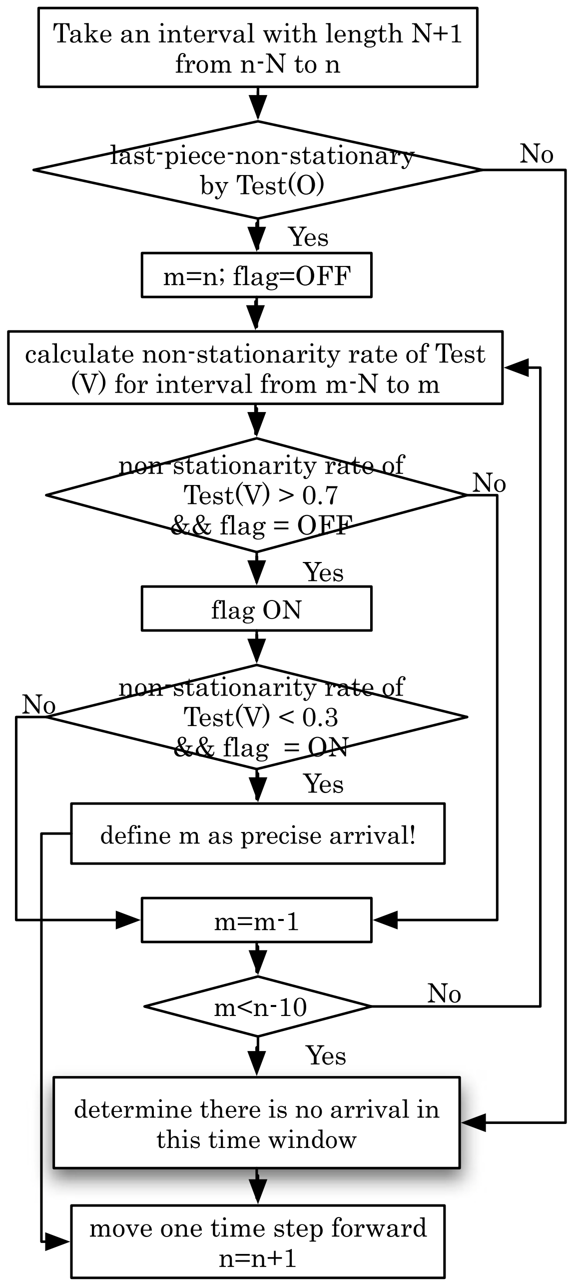


Fig. 5. The outline of the algorithm of our method for automatic picking.

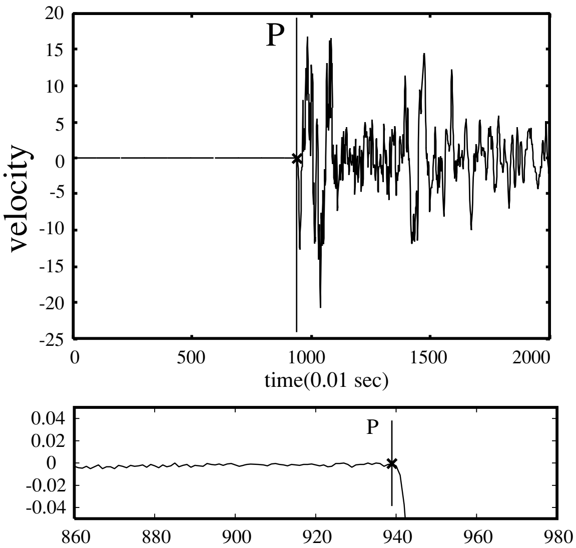
4. Results

We have tested the Test(S) picker using the velocity waveform data set of events selected from National Research Institute for Earth Science and Disaster Prevention's high-sensitivity seismograph network (Hi-net) Japan data from February, 2002 to July, 2003. Selected events have a magnitude ranging from M 0.3 to M 5.5 and epicentral distances less than 30 km from the stations. We used the up-down component (1D) for the P-wave determination and the horizontal component (2D) for the S-wave determination. We tested the P-wave picker on 334 seismograms and the S-wave picker on 117 seismograms.

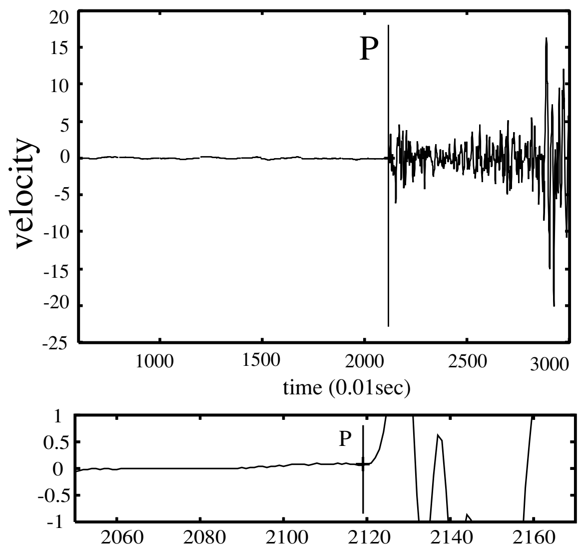
Some examples of the data to which our picker was applied are shown in Fig. 6. In Table 1, we show the ratio of the points picked by our picker within any time lag between the point picked by our picker and the point determined manually by human experts. 90% of the picked points by our picker are within 0.1 s of the manual picks. For comparison, we also applied the AR-model picker (auto_pick

Table 1. Time lags of the point picked as *P*-wave arrival by automatic picker from the point picked by manually picking. Results of our picker and of the picker using the AR-model are both shown. Number of seismograms that could be picked by each method is also showed. The total number of seismograms used to test both methods is 334.

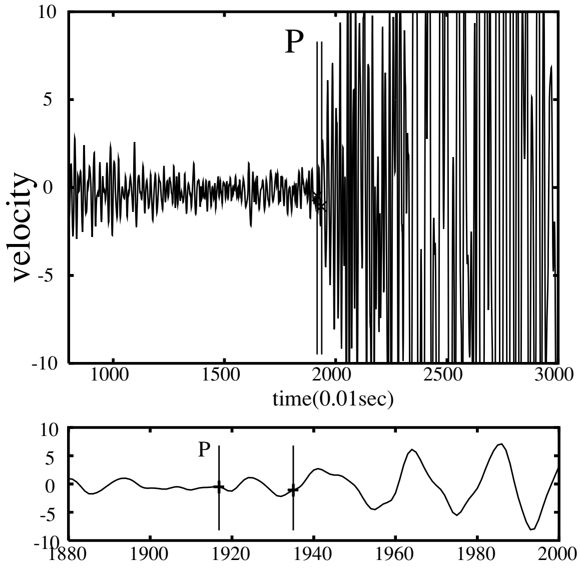
	Number of picked seismograms	Time lag from manually picked point(%)				
		< 0.1 s	0.1–0.5 s	0.5–1 s	1–5 s	>5 s
Our method	278	90.3	9.7	0.0	0.0	0.0
AR picker	290	89.3	1.4	0.3	1.4	7.6



(a)



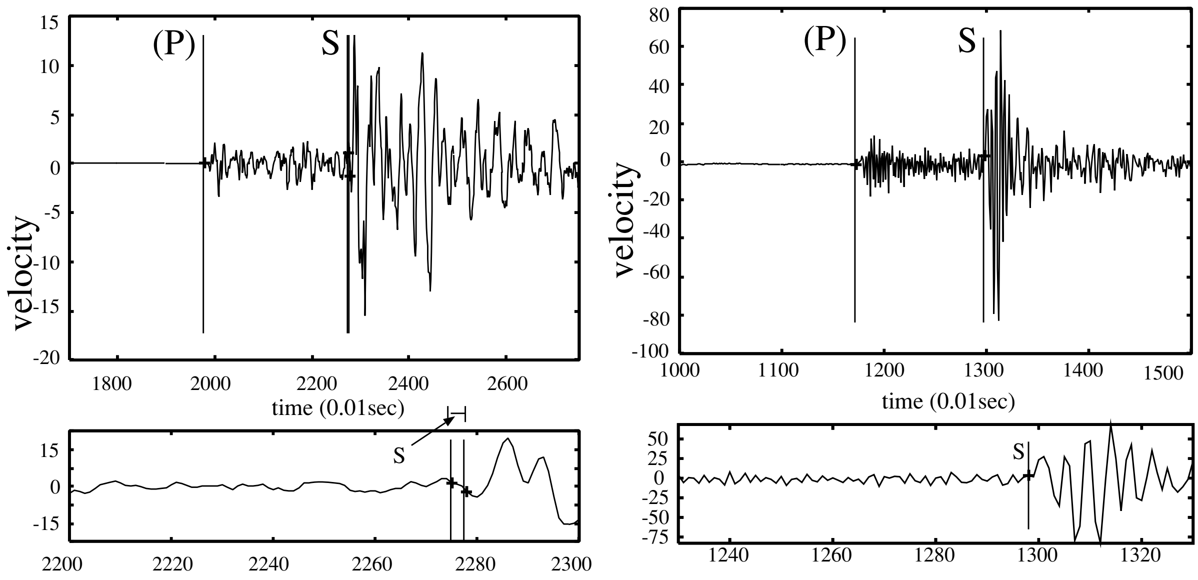
(b)



(c)

Fig. 6. The results of our *P*-wave picker. Seismograms include the *P*-wave arrivals, and the vertical lines on the seismograms indicate the estimated *P*-wave arrivals picked by our method of stationary analysis

program from the Win system (Yokota *et al*., 1981), a major seismic monitoring system in Japan) to the same data set. For the AR-model picker of Win system, both the forward and backward AR models (Takanami and Kitagawa, 1988; Leonard, 2000) are used to derive the AIC. The result of the AR-model picker is also shown in Table 1. In the case that the time lag is within 0.1 s, the ratio of the points picked by AR-model picker is almost the same as that of our picker. However, in the case that the time lag is bigger than 0.1 s, some points picked by AR-model picker are determined largely apart from the arrival points estimated by manually picking. These mistakes in AR-model picker are mainly caused by the failure of taking ranges prior to the appl1cation of the picker. The AR-model picker deter-



(a) (b)

Fig. 7. The results of our S-wave picker. Seismograms include the S-wave arrivals and the vertical lines on the seismograms indicate the estimated S-wave arrivals picked by our method of stationary analysis.

mines the point of arrival by (1) determination of the dividing point, (2) estimation of the AR-coefficient from (some part of or whole of) the data before (and after) the dividing point, (3) derivation of the AIC value from the prediction error of the estimated forward (and backward) model, (4) move the dividing point along the time line and the point for the minimum value of AIC is defined as the phase arrival. So, we must take a range including the P-wave arrival prior to applying the AR-model picker by using some other algorithms, for example the STA/LTA ratio. If the range is not taken correctly, AR-model picker fails to pick P-wave arrival. And even if we could take the range correctly, the AIC value sometimes has several peaks, and then the AR-model picker may also determine the wrong point as the arrival point by picking the wrong peak. Our picker takes intervals along the time line, and does not have to take a prior range. This difference may be significant especially for an earthquake with relatively low signal/noise ratio; for example, volcanic earthquakes. This property is also useful when the picker is used for real time monitoring. Overall, our picker picked the P-wave arrival with considerable accuracy, and it was able to pick the seismograms which AR-model picker could not pick.

We constructed the S-wave picker on the basis of the same algorithm as the P-wave picker. We used the 2-D horizontal component of seismograms for the S-wave picker. A total of 117 seismograms with the manually picked S-wave arrivals are used to test the algorithm, and 108 of these could be determined by the S-wave arrival automatically by our picker. Some of the examples are shown in Fig. 7. For the comparison to the P-wave picker, we use the same time length $N+1=100$ for the S-wave picker. The S-wave picker is constructed for the 2-D data. Then, M turns into 9. This value is relatively smaller than the case of the P-wave picker. However, we have tested the S-wave picker with longer time-length toward some seismograms and obtained a similar result as the case of $M=9$. It would also be useful to use a longer time length, let's say, more than 200 for the application of 2-D or multidimensional data case, with the consideration of the computational time. S-wave picking is more difficult and less accurate than P-wave picking, as the other picker, because the non-stationarity rate of Test(V) becomes high after P-wave arrival. After the P-wave arrival, many phases of seismic signals, such as surface waves, arrive, and our method also picks these points of arrival, instead of the other method of auto-picking. Improvement needs to be made at this point, but in the present study we used the point nearest to the maximum amplitude in certain length of time from the P-wave arrival as the picked S-wave arrival. The P-wave arrival is of course picked by the method of picking S-wave arrival, and this P-wave arrival is also shown in Fig. 7. However, the estimated P-wave arrival by the S-wave picker is not reliable since the S-wave picker does not use the up-down component of seismograms. A total of 108 seismograms could be picked for the S-wave arrival by our method, and 71.3% of the auto-picks are within 0.1 s of the manual picks. In the case of the AR-model picker, 76 seismograms could be determined as the S-wave arrival and 67% of the AR-model picks are within the range of 0.1 s from the manual picks.

From these applications, our picker is a reliable method to pick the point of arrival automatically. Since, unlike the method of the AR-model picker, our method picks all points that match the criteria, it will be useful in a swarm or for picking many kinds of seismic phases in one event at one time, for example, P-wave, S-wave, surface wave, and converted wave.

5. Conclusions

We have constructed a new detector and picker of seismic wave arrivals based on stationary analysis according to the theory of the KM_2O-Langevin equations. This method does not require the prior assumption for the time series and does not have to take a range preliminarily before applying picker. We applied this method to real seismic data and confirmed that our method is accurate enough to use as a part of the seismic early detection system and that it can pick P-wave arrival which could not be picked by method using AR-model.

Acknowledgments. We are grateful to Dr. T. Takanami for beneficial discussions and comments. We are also thankful to Dr. S. Kaneshima for helpful comments and suggestions. Dr. M. Leonard's advices are helpful for us and we appreciate about that. This research was supported by Grant-in-Aids for International Scientific Research (B) (No. 14340127) and (No. 14340030) from the Ministry of Education, Culture, Sports, Science and Technology, Japan, and Super robust Computation Project of the 21st Century COE Program "Information Science and Technology Strategic Core".

Appendix A.

If **X** is a local and weakly stationary process, special relationships exist among the elements of $LM(\mathbf{X})$. The Characterization Theorem describes these relationships for Stationary Property (Okabe, 1999).

Theorem 1 (Characterization Theorem for Stationary Property)

The necessary and sufficient condition for **X** *to have a weakly stationary property is that* $LM(\mathbf{X})$ *satisfies the Dissipation-Dissipation Theorem (DDT) and Fluctuation-Dissipation Theorem (FDT).*

Dissipation-Dissipation Theorem(DDT)

For each integer $(1 \le n \le N)$

$$\gamma_\pm(n, 0) = \delta_\pm(n) \tag{A.1}$$

For each integer $(1 \le k < n \le N)$

$$\gamma_\pm(n,\ k) = \gamma_\pm(n-1, k-1) + \delta_\pm(n)\gamma_\mp(n-1, n-k-1) \tag{A.2}$$

Fluctuation-Dissipation Theorem(FDT)

For each integer $(1 \le n \le N)$

$$\begin{aligned} V_+(0) &= V_-(0) \\ V_\pm(n) &= (\mathbf{I} - \delta_\pm(n)\delta_\mp(n))V_\pm(n-1) \\ \delta_+(n)V_-(n-1) &= V_+(n-1)\,{}^t\delta_-(n) \\ \delta_+(n)V_-(n) &= V_+(n)\,{}^t\delta_-(n) \end{aligned} \tag{A.3}$$

On the other hand, we can uniquely construct the KM$_2$O-Langevin matrix system which satisfies the (DDT) and (FDT) from any positive definite d × d matrix function with the Toeplitz condition (Okabe, 2000). This theorem is called the Construction Theorem.

Theorem 2 (Construction Theorem)

For any matrix function $\mathbf{R} = (R(n); |n| \le N)$ *which satisfies the Toeplitz condition and* ${}^tR(n) = R(-n)$ $(0 \le n \le N)$*, there exists a unique system* $LM(\mathbf{R})$ *of* $d \times d$ *matrices which satisfies (DDT),(FDT) and the following relationship (PAC).*

$$\text{(PAC)}\quad \delta^R_\pm(n+1) = -(R(\pm(n+1)) + \sum_{k=0}^{n-1}\gamma^R_\pm(n,k)R(\pm(k+1)))V^R_\mp(n)^{-1} \tag{A.4}$$

On the basis of Theorem 2, $LM(\mathbf{R})$ can be calculated by the following algorithm.

[Step 1] $V^R_\pm(0) = R(0)$

[Step 2] $\delta^R_\pm(1) = -R(\pm 1)V^R_\mp(0)^{-1}$

[Step 3] $\gamma^R_\pm(1,0) = \delta^R_\pm(1)$

[Step 4] $V^R_\pm(1) = (I - \delta^R_\pm(1)\delta^R_\mp(1))V^R_\pm(0)$

Assuming that we could obtain $\{\gamma^R_\pm(m,k),\ \delta^R_\pm(m), V^R_\pm(l);\ 0 \le k < m \le n,\ 0 \le l \le n\}$ for n of $1 \le n \le N-1$, then we can calculate terms of $n+1$.

[Step 5]

$$\delta^R_\pm(n+1) = -\ (R(\pm(n+1)) + \sum_{k=0}^{n-1}\gamma^R_\pm(n,k)R(\pm(k+1)))V^R_\mp(n)^{-1}$$

[Step 6]

$$\begin{aligned} \gamma^R_\pm(n+1,0) &= \delta^R_\pm(n+1) \\ \gamma^R_\pm(n+1,k) &= \gamma^R_\pm(n,k-1) + \delta^R_\pm(n+1)\gamma^R_\mp(n,n-k), \quad (1 \le k \le n) \end{aligned}$$

[Step 7]

$$V^R_\pm(n+1) = (I - \delta^R_\pm(n+1)\delta^R_\mp(n+1))V^R_\pm(n)$$

$LM(\mathbf{R})$can be obtained by the recursive execution of this algorithm.

By solving the KM$_2$O-Langevin equations with elements of $LM(\mathbf{R})$as the coefficient, we can make a d-dimensional stochastic process **X**. Then **X** has following property.

Theorem 3

X *has a stationary property and its covariance matrix function is* R*. KM$_2$O-Langevin matrix* $LM(\mathbf{X})$*is equal to* $LM(\mathbf{R})$.

We consider the case that we have any d-dimensional-valued stochastic process **X**, and any non-negative definite matrix function **R** which satisfies the Toeplitz condition and ${}^tR(n) = R(-n)$ $(0 \le n \le N)$. Note that **R** is independent of **X**. From Theorem 2, we obtain the pseudo-KM$_2$O-Langevin matrix from **R**. Using elements of $LM(\mathbf{R})$, we introduce a pseudo-fluctuation process of **X** as follows: $\tilde{\nu}(n) = X(n) + \sum_{k=0}^{n-1}\gamma^R_+(n,k)X(k)$. Considering Theorem 1, 2, and 3, the necessary and sufficient condition for **X** to have a weakly stationary property and to have the covariance matrix function **R** is that $\tilde{\nu}(n)$ is orthogonal, i.e.

$$(\tilde{\nu}(m), \tilde{\nu}(n)) = \delta_{nm}V^R_+(n) \tag{A.5}$$

Appendix B.

In the derivation of the formulated criterion [Orthogonality], a basic concept of the formulation is to transform $R^{(\xi^{(s)})}(n;m)$ into a form that converge to the normal distribution $N(0,1)$. To achieve this purpose, we express $d(M+1)R^{\xi^{(s)}}(n;m)$ as the sum of $R_1^{\xi^{(s)}}(n;m)$ and

$R_2^{\xi^{(s)}}(n;m)$ that are the sum of independent stochastic variables:

$$d(M+1)R^{\xi^{(s)}}(n;m) = R_1^{\xi^{(s)}}(n;m) + R_2^{\xi^{(s)}}(n;m) \quad \text{(B.1)}$$

To obtain this separation, we divide both $d(M+1)$ and m by $2n$ and call the quotients q, s and the reminders r, t, respectively.

$$\begin{array}{l} d(M+1)-1 = q(2n)+r \quad (0 \le r \le 2n-1) \\ m - s(2n) + t \quad (0 \le t \le 2n-1) \end{array} \quad \text{(B.2)}$$

If $0 \le r \le n-1$,

$$R_1^{\xi^{(s)}}(n;m) = \begin{cases} \sum_{k=0}^{n-t-1} \xi^{(s)}(m+k)\xi^{(s)}(m+n+k) \\ \quad + \sum_{j=s+1}^{q-1} (\sum_{k=0}^{n-1} \xi^{(s)}(2jn+k)\xi^{(s)}((2j+1)n+k)) \\ \quad (0 \le t \le n-1) \\ \sum_{j=s+1}^{q-1} (\sum_{k=0}^{n-1} \xi^{(s)}(2jn+k)\xi^{(s)}((2j+1)n+k)) \\ \quad (n \le t \le 2n-1) \end{cases} \quad \text{(B.3)}$$

$$R_2^{\xi^{(s)}}(n;m) = \begin{cases} \sum_{j=s}^{q-2} (\sum_{k=0}^{n-1} \xi^{(s)}((2j+1)n+k)\xi^{(s)}((2j+1)n+k)) \\ \quad + \sum_{k=0}^{r} \xi^{(s)}((2q-1)n+k)\xi^{(s)}(2qn+k) \\ \quad (0 \le t \le n-1) \\ \sum_{k=0}^{n-t-1} \xi^{(s)}(m+k)\xi^{(s)}(m+n+k) \\ \quad + \sum_{j=s}^{q-2} (\sum_{k=0}^{n-1} \xi^{(s)}((2j+1)n+k)\xi^{(s)}((2j+1)n+k)) \\ \quad + \sum_{k=0}^{r} \xi^{(s)}((2q-1)n+k)\xi^{(s)}(2qn+k) \\ \quad (n \le t \le 2n-1) \end{cases} \quad \text{(B.4)}$$

If $n \le r \le 2n-1$

$$R_1^{\xi^{(s)}}(n;m) = \begin{cases} \sum_{k=0}^{n-t-1} \xi^{(s)}(m+k)\xi^{(s)}(m+n+k) \\ \quad + \sum_{j=s+1}^{q-1} (\sum_{k=0}^{n-1} \xi^{(s)}(2jn+k)\xi^{(s)}((2j+1)n+k)) \\ \quad + \sum_{k=0}^{r-n} \xi^{(s)}(2qn+k)\xi^{(s)}((2q+1)n+k) \\ \quad (0 \le t \le n-1) \\ \sum_{j=s+1}^{q-1} (\sum_{k=0}^{n-1} \xi^{(s)}(2jn+k)\xi^{(s)}((2j+1)n+k)) \\ \quad + \sum_{k=0}^{r-n} \xi^{(s)}(2qn+k)\xi^{(s)}((2q+1)n+k) \\ \quad (n \le t \le 2n-1) \end{cases} \quad \text{(B.5)}$$

$$R_2^{\xi^{(s)}}(n;m) = \begin{cases} \sum_{k=0}^{2n-1-t} \xi^{(s)}(m+k)\xi^{(s)}(m+n+k) \\ \quad + \sum_{j=s+1}^{q-1} (\sum_{k=0}^{n-1} \xi^{(s)}((2j+1)n+k)\xi^{(s)}((2j+1)n+k)) \\ \quad (0 \le t \le n-1) \\ \sum_{j=s}^{q-1} (\sum_{k=0}^{n-1} \xi^{(s)}((2j+1)n+k)\xi^{(s)}((2j+1)n+k)) \\ \quad (n \le t \le 2n-1) \end{cases} \quad \text{(B.6)}$$

$L_{n,m}^{(1)}$ and $L_{n,m}^{(2)}$ are defined to be the total number of the term number of $R_1^{\xi^{(s)}}(n;m)$ and $R_2^{\xi^{(s)}}(n;m)$, respectively. From (B1), (B2),

$$L_{n,m}^{(1)} + L_{n,m}^{(2)} = d(M+1) - n - m \quad \text{(B.7)}$$

is satisfied. In concrete terms, these are given as follows:
If $0 \le r \le n-1$,

$$\begin{cases} L_{n,m}^{(1)} = \begin{cases} n(q+s)-m & (0 \le t \le n-1) \\ n(q-s-1) & (n \le t \le 2n-1) \end{cases} \\ L_{n,m}^{(2)} = \begin{cases} n(q-s-1)+r+1 & (0 \le t \le n-1) \\ n(q+s)+r+1-m & (n \le t \le 2n-1) \end{cases} \end{cases} \quad \text{(B.8)}$$

If $n+1 \le r \le 2n-1$

$$\begin{cases} L_{n,m}^{(1)} = \begin{cases} n(q+s-1)+r+1-m & (0 \le t \le n-1) \\ n(q-s-2)+r+1 & (n \le t \le 2n-1) \end{cases} \\ L_{n,m}^{(2)} = \begin{cases} n(q-s) & (0 \le t \le n-1) \\ n(q+s+1)-m & (n \le t \le 2n-1) \end{cases} \end{cases} \quad \text{(B.9)}$$

From the central limit theorem, for each n, m $(1 \le n \le L,\ 0 \le m \le L-n)$, $\left(\sqrt{L_{n,m}^{(1)}}\right)^{-1} R_1^{\xi^{(s)}}(n;m)$, $\left(\sqrt{L_{n,m}^{(2)}}\right)^{-1} R_2^{\xi^{(s)}}(n;m)$ will approximately be the occurrence of a stochastic variable which following the Gaussian normal distribution $N(0,1)$ for sufficiently large M. Therefore,

$$\begin{array}{l} \left(\sqrt{L_{n,m}^{(1)}}\right)^{-1} |R_1^{\xi^{(s)}}(n;m)| < 1.96 \\ \left(\sqrt{L_{n,m}^{(1)}}\right)^{-1} |R_1^{\xi^{(s)}}(n;m)| < 1.96 \end{array} \quad \text{(B.10)}$$

are satisfied in probability of 0.95. From this,

$$d(M+1)\left|R^{\xi^{(s)}}(n;m)\right| = \left|\sqrt{L_{n,m}^{(1)}}\left(\left(\sqrt{L_{n,m}^{(1)}}\right)^{-1} R_1^{\xi^{(s)}}(n;m)\right) + \sqrt{L_{n,m}^{(2)}}\left(\left(\sqrt{L_{n,m}^{(2)}}\right)^{-1} R_2^{\xi^{(s)}}(n;m)\right)\right| < 1.96\left(\sqrt{L_{n,m}^{(1)}} + \sqrt{L_{n,m}^{(2)}}\right) \quad \text{(B.11)}$$

By transforming (B11), we get inequality (C-O).

$$d(M+1)\left(\sqrt{L_{n,m}^{(1)}} + \sqrt{L_{n,m}^{(2)}}\right)^{-1}\left|R^{\xi_i}(n;m)\right| < 1.96 \quad \text{(B.12)}$$

References

Akaike, H. and T. Nakagawa, *Statistical Analysis and Control of Dynamic Systems*, KTK Scientific Publishers, Tokyo, 1988.

Allen, R., Automatic phase pickers: Their present use and future prospects, *Bull. Seism. Soc. Am.*, **72**, 225–242, 1982.

Anant, S. K. and F. U. Dowla, Wavelet transform methods for phase identification in three-component seismogram, *Bull. Seism. Soc. Am.*, **87**, 1598–1612, 1997.

Leonard, M., Comparison of manual and automatic onset time picking, *Bull. Seism. Soc. Am.*, **90**, 1384–1390, 2000.

Leonard, M. and B. L. N. Kennett, Multi-component autoregressive techniques for the analysis of seismograms, *Phys. Earth Planet. Interiors*, **113**, 247–264, 1999.

Maeda, N., A method for reading and checking phase times in auto-processing system of seismic wave data, *Zisin*, **38**, 365–379, 1985.

Matsuura, M. and Y. Okabe, On a non-linear prediction problem for one-dimensional stochastic processes, *Japan J. Math.*, **27**, 51–112, 2001.

Okabe, Y., On the theory of KM_2O-Langevin equations for stationary flows (1): characterization theorem, *J. Math. Soc. Japan.*, **51**, 817–841, 1999.

Okabe, Y., On the theory of KM_2O-Langevin equations for stationary flows (2): construction theorem, *Acta Applicandae Mathematicae.*, **63**, 307–322, 2000.

Okabe, Y. and Y. Nakano, The theory of KM_2O-Langevin equations and its applications to data analysis (I): stationary analysis, *Hokkaido Math. J.*, **20**, 45–90, 1991.

Okabe, Y. and T. Yamane, The theory of KM_2O-Langevin equations and its applications to data analysis (II): deterministic analysis, *Nagoya Math. J.*, **152**, 175–201, 1998.

Okabe, Y., M. Matsuura, M. Takeo, and H. Ueda, On an abnormality test for detecting initial phases of earthquakes, *Math. Eng. Tech. Rep.*, **METR-2003-41**, Department of Mathematical Informatics, The University of Tokyo, 2003.

Reading, A. M., W. Mao, and D. Gubbins, Polarization filtering for automatic picking of seismic data and improved converted phase detection, *Geophys. J. Int.*, **147**, 227–234, 2001.

Sleeman, R. and T. van Eck, Robust automatic P-phase picking: an on-line implementation in the analysis of broadband seismogram recordings, *Phys. Earth Planet. Interiors*, **113**, 265–275, 1999.

Snedecor, G. W., W. G. Cochran, ISU Statistics Depts. Staff, D. F. Cox, *Statistical Methods*, 8th ed., pp. 53–55, pp. 71–73, Blackwell Publishing Limited, 1989.

Student (Gosset, W. S.), The probable error of a mean, *Biometrika*, **6**, 1., 1–25, 1908

Takanami, T. and G. Kitagawa, A new efficient procedure for the estimation of onset times of seismic waves, *J. Phys. Earth*, **36**, 267–290, 1988.

Vidale, J. E., Complex polarization analysis of particle motion, *Bull. Seism. Soc. Am.*, **76**, 1393–1405, 1986.

Withers, M., R. Aster, C. Young, J. Beiriger, M. Harris, S. Moore, and J. Trujikko, A comparison of select trigger algorithms for automated global seismic phase and event detection, *Bull. Seism. Soc. Am.*, **88**, 95–106, 1998.

Yokota, T., S. Zhou, M. Mizoue, and I. Nakamura, An automatic measurement of the arrival time of seismic waves and its application to an on-line processing system, *Bull. Earthq. Res. Inst. Univ. Tokyo*, **55**, 449–484, 1981 (in Japanese with English abstract).

Zhang, H., C. Thurber, and C. Rowe, Automatic P-Wave arrival detection and picking with multiscale wavelet analysis for single-component recordings, *Bull. Seism. Soc. Am.*, **93**, 1904–1912, 2003.

S. Nakamula (e-mail: show@eri.u-tokyo.ac.jp), M. Takeo, Y. Okabe, and M. Matsuura