Phase Transition Model of Earthquakes

Hiroyuki Kikuchi[1]

[1]Seismic Lab

November 22, 2022

Abstract

Instrumentally measurable earthquake precursors are derived by considering earthquakes as a phase transition in cellular automaton. The existence of the phase transition in CA-184 is implied by considering the space-time diagram of the CA-184 as the worldsheet of the Polyakov action in string theory. The CA-184 is the rule-184 cellular automaton (CA), which is a special case of Burgers cellular automaton (BCA) rigorously derived through transforming the one-dimensional Burgers equation. Then, p-CA-184, the CA-184 with probabilistic fluctuation and p-BCA, the BCA with probabilistic fluctuation, are associat-ed with the earthquake. The Fourier transforms of p-CA-184 and p-BCA dynamics near the phase transition reproduce the Fourier transforms of the ground vibration data before and after the earthquake, respectively. Consequently, we consider an earthquake as the phase transition of CA-184. Two precursors of the phase transition of the CA-184, therefore the earthquake precursors, are derived with p-CA-184 by introducing the Gumbel distribution defined in the framework of extreme value theory. To evaluate the pre-cursors, the ground vibration data measured at three locations over a period of approxi-mately 10 years has been investigated. One of the derived precursors is observed before every studied earthquake with seismic intensity greater than 4, and the other precursor is observed selectively before the large earthquake of magnitude 9. Furthermore, the two pre-cursors calculated for different frequencies and time scales are observed at similar timing before the magnitude 9 earthquake. The phase transition model of earthquakes provides the practical and reliable earthquake prediction method.

Phase Transition Model of Earthquakes

Hiroyuki Kikuchi[1*]

[1]Seismic Lab, 307, 12-1, 2-Chome, Wakamiya, Okegawa-shi, Saitama 363-0022, Japan

Key Points:

- Existence of the phase transition in the periodic rule 184 cellular automaton is implied by string theory
- Periodic rule 184 cellular automaton with stochastic fluctuations near the phase transition reproduces ground vibrations before earthquakes
- Precursors of the phase transition of the periodic rule 184 cellular automaton are earthquake precursors

*307, 12-1, 2-Chome, Wakamiya, Okegawa-shi, Saitama 363-0022, Japan

Corresponding author: Hiroyuki Kikuchi, hkikuchi@kkh.biglobe.ne.jp

Abstract

Instrumentally measurable earthquake precursors are derived by considering earthquakes as a phase transition in cellular automaton. The existence of the phase transition in CA-184 is implied by considering the space-time diagram of the CA-184 as the worldsheet of the Polyakov action in string theory. The CA-184 is the rule-184 cellular automaton (CA), which is a special case of Burgers cellular automaton (BCA) rigorously derived through transforming the one-dimensional Burgers equation. Then, p-CA-184, the CA-184 with probabilistic fluctuation and p-BCA, the BCA with probabilistic fluctuation, are associated with the earthquake. The Fourier transforms of p-CA-184 and p-BCA dynamics near the phase transition reproduce the Fourier transforms of the ground vibration data before and after the earthquake, respectively. Consequently, we consider an earthquake as the phase transition of CA-184. Two precursors of the phase transition of the CA-184, therefore the earthquake precursors, are derived with p-CA-184 by introducing the Gumbel distribution defined in the framework of extreme value theory. To evaluate the precursors, the ground vibration data measured at three locations over a period of approximately 10 years has been investigated. One of the derived precursors is observed before every studied earthquake with seismic intensity greater than 4, and the other precursor is observed selectively before the large earthquake of magnitude 9. Furthermore, the two precursors calculated for different frequencies and time scales are observed at similar timing before the magnitude 9 earthquake. The phase transition model of earthquakes provides the practical and reliable earthquake prediction method.

1 Introduction

The earthquake prediction with the specified time, location and magnitude has been studied for over 100 years without successfully generating the publicly acceptable prediction method (Geller, 1997). The difficulty in the theoretical earthquake model constituting springs, mass and friction is that no practically effective measurement methods exist for acquiring the friction data which are complex and heterogeneously distributed on the earth. The empirical earthquake predictions are believed to be untrustworthy since many of the claimed precursors are observed after earthquakes and since they do not link to physical reasoning. There is a negative opinion that the earthquake prediction is intrinsically impossible since the earth is in a state of self-organized criticality, in which any small earthquake possibly results in a large earthquake through cascading.

In this paper, we propose the theoretical earthquake model, which involves no mechanical parameters such as spring, mass or friction, for deriving the measurable earthquake precursors, instead of building the model for simulating the earthquake itself. The rest of this paper is organized as follows: We firstly review the mathematical formalism, on the Burgers cellular automaton (BCA) obtained by transforming the one-dimensional Burgers equation, its special case CA-184, which is the rule-184 cellular automaton (CA), and their stochastic extensions. Secondly, the existence of the phase transition in CA-184 is shown. Thirdly, p-CA-184 and p-BCA, which are respectively CA-184 with probabilistic fluctuation and BCA with probabilistic fluctuation, have been related to the ground vibration near earthquakes. Then, we consider an earthquake as the phase transition of the CA-184. Fourthly, the precursor of the phase transition, therefore of the earthquakes, is derived by introducing Gumbel distribution applied to the output of the p-CA-184 calculation. Finally, in order to confirm the reliability of the derived earthquake precursors, the precursors have been searched for in the ground vibration data measured at three locations over a period of approximately 10 years.

2 Mathematical formalism

The CA-184 and the BCA are briefly reviewed at first. Secondly, p-CA-184 and p-BCA which are the probabilistic extension of CA-184 and BCA are defined. Then, in order to discuss the motion and the entropy of CA-184, Nambu-Goto action and Polyakov action in string theory are introduced. In discussing string theory, we focus on the formalism of the variational principle, rather than on the physical contents. The variational principle states that the object moves in the way that the action (S) is minimized, and therefore in the way that the variational of the action becomes zero ($\delta S = 0$). The equation of motion is resulted as the condition for the $\delta S = 0$. In the review of the Conformal Field Theory (CFT), the central charge is in presence in the conformal transformation of the stress-energy tensor which is the variational of the action with respect to the metric. The relation between central charge and entropy is explained at the end of this section.

2.1 CA-184 and BCA (Nishinari & Takahashi, 1999)

BCA which includes CA-184 has been derived from Burgers equation through difference Burgers equation. The results are described below:

$$\frac{dv}{dt} = 2v\frac{dv}{dx} + \frac{d^2v}{dx^2} \qquad \textit{Burgers equation} \tag{1}$$

where v, t and x denote real variable, time and space.

$$u_j^{t+1} = u_{j-1}^t + \frac{1 + \frac{1-2\delta}{c\delta}u_j^t + \frac{1}{c^2}u_j^t u_{j+1}^t}{1 + \frac{1-2\delta}{c\delta}u_{j-1}^t + \frac{1}{c^2}u_{j-1}^t u_j^t} \qquad \textit{Difference Burgers equation} \tag{2}$$

where $\delta = \Delta t/(\Delta x)^2$. The natural numbers t and x denote respectively time and space, and c is a real constant. Substituting $u_j^t/c = e^{\Delta x v} \simeq 1 + \Delta x v + (\Delta x v)^2/2$, we obtain Eq. (1) from Eq. (2).

BCA is obtained by discretizing the dependent variable u in Eq. 2, which is called 'ultradiscretization':

$$U_j^{t+1} = U_j^t + min(\overline{M},\, U_{j-1}^t,\, \overline{L} - U_j^t) - min(\overline{M},\, U_j^t,\, \overline{L} - U_{j+1}^t) \qquad BCA \tag{3}$$

where U, $\overline{M}$, $and\,\overline{L}$ are integer. In the case that $\overline{L} = 1 \leq \overline{M}$, CA-184 is obtained:

$$U_j^{t+1} = U_j^t + min(U_{j-1}^t, 1 - U_j^t) - min(U_j^t, 1 - U_{j+1}^t) \qquad CA - 184 \tag{4}$$

CA-184 which is frequently introduced for modeling traffic jams is the one-dimensional array of cells, each containing 0 or 1. The time evolution of the cells, which depends only on the cells on both side of it, follows the rule-184:

$$\frac{U_{j-1}^t U_j^t U_{j+1}^t}{U_j^{t+1}} = \frac{111}{1}, \frac{110}{0}, \frac{101}{1}, \frac{100}{1}, \frac{011}{1}, \frac{010}{0}, \frac{001}{0}, \frac{000}{0} \tag{5}$$

where t and j are natural numbers, indicating time and space, respectively. The denominator U_j^{t+1}, the value of the cell at the (time, space) = $(t+1, j)$, is determined by the numerator $U_{j-1}^t U_j^t U_{j+1}^t$, the sequence of three binary numbers at time t. The rule-184 states that when there is the sequence 10 at time t, it will become 01 at time t+1. In the other words, assuming 0 and 1 denote empty cell and occupied cell respectively, if there is an empty cell on the right side at t, the empty cell will be occupied at t+1. Therefore, the occupied cells move to the right as time passes. There are eight possible time evolution patterns as shown in the righthand side of Eq. (5). The binary number 10111000 obtained by arranging the denominators in order is 184 in decimal notation, which is the origin of the name rule-184.

2.2 *p*-CA-184 and *p*-BCA : CAs with probabilistic fluctuations

Probabilistic fluctuation needs to be introduced in CA for expressing the complex phenomena such as ground vibration on the earth. Introducing the moving probability p, the rule-184 (Eq. (5)) can be rephrased that if there is a sequence 10 in the numerator at time t, then with probability $p = 1$, the sequence become 01 at time t+1. Extending the probability p to an arbitrary real number, p-CA-184 is defined as the following: if there is a sequence 10 in the numerator at time t, then with probability p, the sequence become 01 at time t+1. In the other words, when p-CA-184 advances, it advances with probability p. The CA-184 with the moving probability p is called Asymmetric Simple Exclusion Process (ASEP). However, ASEP is referred to as p-CA-184 in this paper since p represents the state in the phase transition process described in the later sections.

With reference to Eq. (5), the advancing conditions for p-CA-184 are:

$$\frac{U_{j-1}^t U_j^t U_{j+1}^t}{U_j^{t+1}} = \frac{110}{0}, \frac{101}{1}, \frac{100}{1}, \frac{010}{0} \tag{6}$$

which is equivalent to

$$U_j^{t+1} \neq U_j^t \tag{7}$$

With reference to this advancing condition as well as Eq. (3) and Eq. (4), the protocols of p-CA-184 and p-BCA are defined as the following:

$$\begin{aligned}
&p - CA - 184\ : \\
&\quad If\, U_j^{t+1} \neq U_j^t\,, then\,, \\
&\qquad U_j^{t+1} = U_j^t + min(U_{j-1}^t, 1 - U_j^t) - min(U_j^t, 1 - U_{j+1}^t) \quad with\, probability = p\,, and \\
&\qquad U_j^{t+1} = U_j^t \quad with\, probability = 1 - p\,.
\end{aligned} \tag{8}$$

$$\begin{aligned}
&p - BCA\ : \\
&\quad If\, U_j^{t+1} \neq U_j^t\,, then\,, \\
&\qquad U_j^{t+1} = U_j^t + min(\overline{M},\, U_{j-1}^t,\, \overline{L} - U_j^t) - min(\overline{M},\, U_j^t,\, \overline{L} - U_{j+1}^t) \quad with\, probability = p\,, and \\
&\qquad U_j^{t+1} = U_j^t \quad with\, probability = 1 - p\,.
\end{aligned} \tag{9}$$

2.3 Namb-Goto and Polyakov action (Polchinski, 1998)

$Nambu\!-\!Goto$ action which is proportional to the area of worldsheet that is the trajectory of an object, and its convenient form, Polyakov action, are described respectively in Eq. (10) and Eq. (11).

$$\begin{aligned}
&Nambu - Goto\,action: \\
&\qquad S_{NG} = -T\int_M d\mathcal{A} \\
&\qquad d\mathcal{A} = d\tau d\sigma(-\gamma)^{1/2} \quad ; \quad \gamma = det\,\gamma_{ab}
\end{aligned} \tag{10}$$

$$\begin{aligned}
&Polyakov\,action: \\
&\qquad S_p = -\frac{T}{2}\int_M d\tau d\sigma(-h)^{1/2}h^{ab}G_{ab} \\
&\qquad G_{ab} = g_{\gamma\lambda}\partial_a X^\gamma \partial_b X^\lambda \\
&\qquad h_{ab} = h_{ab}(\tau,\sigma) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad , \quad g_{\nu\mu} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\
&\qquad h_{ab}(-h)^{-1/2} = \gamma_{ab}(-\gamma)^{-1/2}
\end{aligned} \tag{11}$$

where T, M, $\mathcal{A}$ denote constant, worldsheet, and area of worldsheet, respectively. γ_{ab} and h_{ab} are worldsheet metrics of which determinants are respectively γ and h. $g_{\nu\mu}$ is the metric on X which specifies the position of a point on a string in two-dimensional space-time . The matrix representation of the metric h_{ab} indicates that the worldsheet is defined in Minkowski space. The subscripts a, b, ν and μ are 0 or 1. The parameters $\xi^1 = \sigma$ and $\xi^0 = \tau$ are indexes respectively on a string and on a trajectory of the string. The string parametrized by σ moves in the two dimensional space-time ($X^0(\tau,\sigma) = time$, $X^1(\tau,\sigma) = space$), and the trajectory of the string forms the worldsheet described with the parameters τ and σ (Fig. 1). The formulae follow the Einstein summation convention. The partial derivative is described as $\partial_a = \partial/\partial x^a$.

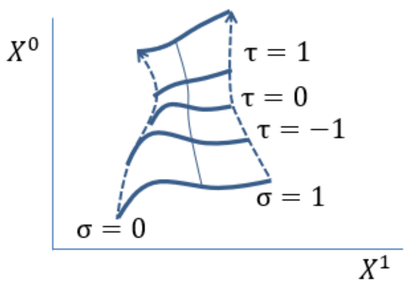


Figure 1: Worldsheet. The string is spanned along σ (solid line) at time τ. The time axis is shown by the dashed lines.

Polyakov action remains constant over changing the parameter τ and σ, which do not have physical meanings. Therefore, the variational of S_p with respect to $X^\gamma(\tau,\sigma)$ is zero, resulting in the equation of motion or Euler-Lagrange equation of the string. The string motion is governed by the wave equation.

$$\begin{aligned}
&\frac{\delta S_p}{\delta X^\gamma} = 0 \quad \Rightarrow \quad \left(\frac{\partial^2}{\partial\tau^2} - \frac{\partial^2}{\partial\sigma^2}\right)X^\gamma = 0 \qquad Wave\,equation \\
&\qquad (\partial_\tau - \partial_\sigma)\,X^\gamma = 0 \qquad Right\,going\,traveling\,wave \\
&\qquad (\partial_\tau + \partial_\sigma)\,X^\gamma = 0 \qquad Left\,going\,traveling\,wave
\end{aligned} \tag{12}$$

where δ denotes variational. The Stress-energy tensor T^{ab} which is involved in discussing entropy is defined as the quantity proportional to the variational of S_p with respect to the metric h_{ab}.

$$\begin{aligned} T^{ab}(\tau,\,\sigma) &\equiv \frac{-4\pi}{\sqrt{-h}}\frac{\delta S_p}{\delta h_{ab}} \qquad\qquad Stress-energy\,tensor \\ T_{ab}(\tau,\,\sigma) &= T\left(G_{ab}-\frac{1}{2}h_{ab}G^c_c\right) = \begin{pmatrix} T_{00} & T_{01} \\ T_{10} & T_{11} \end{pmatrix} \end{aligned} \tag{13}$$

where $T_{00} = T_{11} = \dot{X}^2 + X'^2{=}(\partial X^\mu/\partial\tau)(\partial X_\mu/\partial\tau) + (\partial X^\mu/\partial\sigma)(\partial X_\mu/\partial\sigma)$ and $T_{01} = T_{10} = 2\dot{X}\cdot X'{=}2(\partial X^\mu/\partial\tau)(\partial X_\mu/\partial\sigma)$.

2.4 Conformal Field Theory (Polchinski, 1998)

In order to discuss entropy associated with Conformal Field Theory (CFT), the wave equation (Eq. (12)) and the stress-energy tensor (Eq. (13)) are transformed from Minkowski space-time $(\tau,\,\sigma)$ to Euclidean space-time $(\sigma^0,\,i\sigma^1)$, then to the two-dimensional complex plane $(\omega,\,\overline{\omega}) = (\sigma^0 + i\sigma^1,\,\sigma^0 - i\sigma^1)$:

$$\partial_{\overline{\omega}}\,\partial_\omega X^\mu = 0 \qquad\qquad Equation\,of\,motion \tag{14}$$

$$\begin{aligned} &\partial_\omega X^\mu\,is\,holomprphic\,(left-moving),\;function\,of\,\omega \\ &\partial_{\overline{\omega}} X^\mu\,is\,antiholomprphic\,(right-moving),\;function\,of\,\overline{\omega} \end{aligned}$$

$$\begin{aligned} T(\omega\,\overline{\omega}) &= \begin{pmatrix} T(\omega) & 0 \\ 0 & T(\overline{\omega}) \end{pmatrix} \qquad\qquad Stress-energy\,tensor \\ T(\omega) &= T_{\omega\,\omega} = \frac{1}{2T}(T_{00}+T_{01}) = 2g_{\mu\nu}\partial_\omega X^\mu \partial_\omega X^\nu \\ T(\overline{\omega}) &= T_{\overline{\omega}\,\overline{\omega}} = \frac{1}{2T}(T_{00}-T_{01}) = 2g_{\mu\nu}\partial_{\overline{\omega}} X^\mu \partial_{\overline{\omega}} X^\nu \end{aligned} \tag{15}$$

where $T_{\omega\,\omega}$ and $T_{\overline{\omega}\,\overline{\omega}}$ are respectively holomprphic and antiholomprphic, from Eq. (14).

Central charge is involved in the conformal transformation of the stress-energy tensor from $(z = exp(\omega),\,\overline{z} = exp(\overline{\omega}))$ space-time to $(\omega,\,\overline{\omega})$ space-time:

$$\begin{aligned} T(\omega) &= \left(\frac{\partial\omega}{\partial z}\right)^{-2}\left[T(z) - \frac{cc}{12}\{\omega;z\}\right] = \left(\frac{\partial\omega}{\partial z}\right)^{-2} T(z) - \frac{cc}{24} \qquad ;\quad cc\,:\,Central\,charge \\ \{\omega;z\} &= \frac{1}{(\partial_z\omega)^2}\left[(\partial_z\omega)\left(\partial_z^3\omega\right) - \frac{3}{2}\left(\partial_z^2\omega\right)^2\right] \qquad ;\,Schwarzian \end{aligned} \tag{16}$$

where $T(\omega)$ and $T(z)$ are respectively $T_{\omega\,\omega}$ and $T_{z\,z}$. The similar transformation with the central charge replaced by $\overline{cc}$ which may not be equal to cc holds for the antiholomorphic functions $T(\overline{\omega})$ and $T(\overline{z})$. The $T_{z\,z}$ does not transform as a tensor since it contains the additional term including the central charge. Direct calculations show that the Schwarzian derivative of any Möbius transformation $g(z) = (a'z + b')/(c'z + d')$ is zero.

$$\left\{\frac{a'z+b'}{c'z+d'};z\right\} = 0 \qquad ;a',\,b',\,c',\,and\,d'\,are\,constant \tag{17}$$

2.5 Entropy of the Conformal Field Theory

The central charge in the CFT is related to the thermodynamic entropy of a black hole (Carlip, 1999):

$$Entropy = log\, \rho_n = 2\pi\sqrt{\frac{cc\, L_0}{6}} \propto \sqrt{cc} \tag{18}$$

where ρ_n is number of states. L_0 and the central charge cc are real positive numbers. L_0 is associated with the eigenstate with eigenvalue zero. L_m is known as Virasoro generator which spans Virasoro algebra (Polchinski, 1998):

$$\begin{aligned} L_m &= \frac{1}{2\pi i}\oint dz\, z^{m+1} T_{zz}(z) \\ [L_m,\, L_n] &= (m-n)L_{m+n} + \frac{cc}{12}(m^3 - m)\delta_{m,\,-n} \qquad Virasoro\, algebra \end{aligned} \tag{19}$$

The equation of the entropy for the antiholomorphic terms is obtained from Eq. (18) by substituting cc with $\overline{cc}$ and L_0 with $\overline{L}_0$ which is associated with $T_{\overline{z}\,\overline{z}}$ and $\overline{z}$ (Carlip, 2000). Here, we avoid arguing whether a black hole is equivalent to an earthquake. Rather, we agree that the CFT is linked to the thermodynamic entropy of the physical system, and that the entropy is proportional to the square root of the central charge.

2.6 Gumbel distribution

The Gumbel distribution $G_0(x')$ (Charras-Garrido & Lezaud, 2013), its linearized form applied to the fitting of the measured data, and the probability density function $g_0(x')$, which is the derivative of the $G_0(x')$, are as follows:

$$\begin{aligned} G_0(x') &= \exp\left(-e^{-x'}\right) \quad Gumbel\, distribution \qquad , \quad x' \,:\, Stochastic\, variable \\ -log\left(-log(G_0(x'))\right) &= \frac{1}{\eta}x' - \frac{\lambda}{\eta} \qquad , \quad \eta\, and\, \lambda \,:\, constants \\ \frac{1}{\eta} &\quad : Gradient \\ -\frac{\lambda}{\eta} &\quad : \quad y - intercept \\ g_0(x') &= \exp\left(-x' - e^{-x'}\right) \qquad Probability\, density\, function \end{aligned} \tag{20}$$

The protocol example for obtaining the Gumbel distribution $G_0(x')$, which is the cumulative distribution, is shown in Table 1. Fill the table one by one from left to right to obtain $G_0(x')$. Once the $G_0(x')$ is obtained, the gradient $1/\eta$, y-intercept $-\lambda/\eta$, and the square of the Pearson's correlation coefficient r (Egghe, L. & Leydesdorff, L., 2009) are obtained through the regression analysis. In Table 1, $Class(j)$, ω_j, and x'_j are respectively the row index, the data interval, and the mean of the two numbers of the ω_j. The number of the interval is assumed to be 40, and all data are assumed to be greater than -2.0 and less than or equal to 2,0. N_j is the number of data assumed to be included in the interval ω_j. N_{sj} is the cumulative sum. $G_0(x')$ is obtained by dividing the N_{sj} by the $(N_{total} + 1)$ that is the total number of data plus one.

Table 1: Cumulative distribution $G_0(x'_j)$. The procedure of calculating the cumulative distribution is explained with the detail example.

Class(j)	ω_j	x'_j	N_j	$\sum_{i=1}^{j} N_i \equiv N_{sj}$	$G_0(x'_j) \equiv N_{sj}/(N_{total}+1)$
1	$(-2.0, -1.9]$	-1.95	2	2	$2/(N_{total}+1)$
2	$(-1.9, -1.8]$	-1.85	3	5	$5/(N_{total}+1)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
40	$(1.9, 2.0]$	1.95	1	N_{total}	$N_{total}/(N_{total}+1)$

3 Existence of phase transition in CA-184

The Lagrangian of CA-184 with periodic boundary conditions is related to the Polyakov action of string theory in which Conformal Field theory is related to the entropy of physical systems. Then, the number of states, of which logarithm is entropy, is counted at the phase transition and its neighborhood, of the CA-184. If there is a discontinuous entropy decrease, which indicates an abrupt ordering, then a phase transition probably exists.

3.1 Periodic CA-184

The CA-184 with ten cells and with periodic boundary condition in space is considered. The boundary conditions are determined by applying the rule-184 to $U_{10}^t U_1^t U_2^t / U_1^{t+1}$ and $U_9^t U_{10}^t U_1^t / U_{10}^{t+1}$. The periodic boundary condition represents the stick-slip motion that is thought to exist between two crustal plates. After large t, the periodicity in space yields the periodicity in time in the CA, so that the CA is represented by the area of 10 time-steps by 10 space-steps with the periodic boundary conditions both in time and space. If we consider the integer 1 as an object, and consider the trajectory of the integer as the worldsheet, the worldsheet is stable in time (Fig. 2).

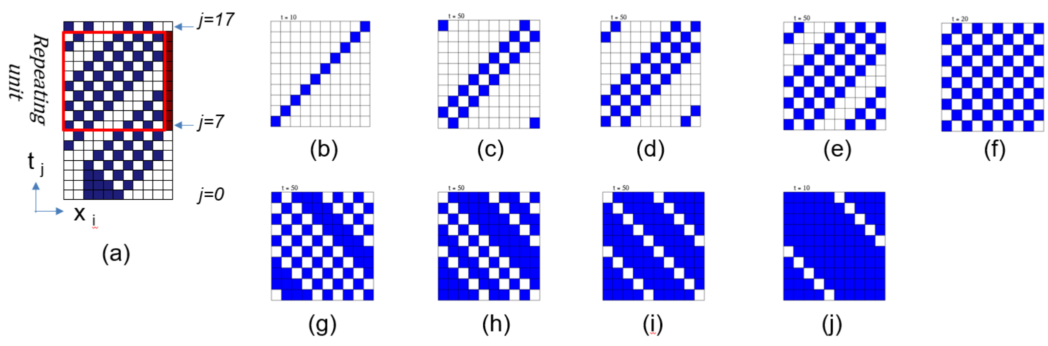


Figure 2: CA-184 worldsheets. The blue cells and the white cells indicate respectively $U_i^j = 1$ and $U_i^j = 0$. t_j and x_i are respectively discretized time and discretized space. ρ is the cell density which is the number of the occupied cell divided by 10, the total number of the cell. (a) Periodicity in time. Since $U_x^{17}{=}U_x^7$, the blue-white pattern between $j = 7$ and $j = 16$ will be repeated after $j = 17$. (b) CA-184 with $\rho = 0.1$. (c) $\rho = 0.2$. (d) $\rho = 0.3$. (e) $\rho = 0.4$. (f) $\rho = 0.5$. (g) $\rho = 0.6$. (h) $\rho = 0.7$. (i) $\rho = 0.8$. (j) $\rho = 0.9$.

3.2 Implication from string theory

If the area of the time stable worldsheet of the CA-184 is considered to be the area of the worldsheet in the Polyakov action in Eq. (11), $\delta S_p = 0$ is guaranteed since the worldsheet is stable. Then, the equation of motion (Eq. (14)), the stress-energy tensor (Eq. (15)), the conformal transformation (Eq. (16)) and the entropy (Eq. (18)) describe the properties of the CA-184. Here, the worldsheet $X^\lambda(\tau, \sigma)$ is assumed to be the continuous representation of the repeating area of the $x-t$ diagram of the CA-184 (Fig. 2). The details of the worldsheet such as the composition points of the strings or the elastic property of the strings are disregarded. It needs to be noted that holomorphic and antiholomorphic functions are independently defined in the four equations. For simplicity in defining the domain of those functions, $\sigma_0 \geq 0$ and $\sigma_1 > 0$ are assumed in $\omega = \sigma_0 + i\sigma_1$ and in $\overline{\omega} = \sigma_0 - i\sigma_1$ so that ω and $\overline{\omega}$ are defined in upper and lower halves of the complex plane, respectively.

From the discussions above, with respect to the CA-184 associated with the Polyakov action, the left-moving motion and the *entropy* ≥ 0 exist in the upper half of the complex plane (ω domain), and the right-moving wave and the *entropy* ≥ 0 exist in the lower half ($\overline{\omega}$ domain).

Now, we paS4y attention to the domain between ω and $\overline{\omega}$ that is the real axis. On the real axis, if we consider the infinitesimal transformation $\omega = \sigma^0 \simeq 0$, then $z = e^\omega \simeq 1+\omega$, or $\omega \simeq z-1$ that is the Möbius transformation of $(a', b', c', d') = (1, -1, 0, 1)$. As it is discussed, for a Möbius transformation, the central charge term becomes zero in Eq. (16), and the stress-energy tensor behaves like a tensor in the conformal transformation from $(z = exp(\omega), \overline{z} = exp(\overline{\omega}))$ to $(\omega, \overline{\omega})$. Consequently, the entropy which is related to the central charge term is zero on the real axis, which implies that greater ordering occurs at the real axis than the surrounding domains. The discontinuous decrease of entropy occurred at the domain between the right-moving domain and the left-moving domain suggests the phase transition.

3.3 S4 symmetry in the periodic CA-184

The checkerboard pattern of the repeating unit of the periodic CA-184 is invariant with respect to the S4 symmetry operation which combines a reflection and a rotation (Fig. 3). In the case of the cell density $\rho = 0.4$ (Fig. 3 (a), (b) and (c)), starting from the initial pattern, the 90 degrees rotation clockwise with respect the origin of the $x_i - t_j$ coordinate system, followed by the reflection with respect to the t_j axis at x_{10} results in the initial pattern. The operation is equivalent to the coordinate transformation $(x,\, t) \to (x',\, t') = (-t,\, -x)$. The Burgers equation (Eq. (1)), which includes CA-184, needs to be satisfied in the $(x',\, t')$ coordinate system.

Consider the Burgers equation (Eq. (21)) in the coordinate system $(x,\, t)$ and the same equation in the $(x',\, t') = (-t,\, -x)$ coordinate system. By direct calculation, the conditions for simultaneously satisfying those two equations are found to be $v_0 = 1/2$ and $v = v(x-2v_0t+c)$, or $v = v(x-t+c)$. The $v = v(x-t+c)$ is the right-going traveling wave since $v = v(0)$ along the line $x-t+c = 0$ or $t = x+c$. This result is consistent with the checkerboard pattern in Fig. 3 (g), where the blue cells align along the dashed line $t_j = x_i + c$ so that the motion of the cell is the right-going traveling wave. From the above discussion and Fig. 2, the motion of the cell is right-going traveling wave not only for the cell density $\rho = 0.4$, but also for $\rho < 0.5$ in which the blue cells align along the line $t_j = x_i + c$.

$$\frac{dv}{dt} = 2v_0\frac{dv}{dx} + \frac{d^2v}{dx^2} \qquad , v_0 \,:\, constant \tag{21}$$

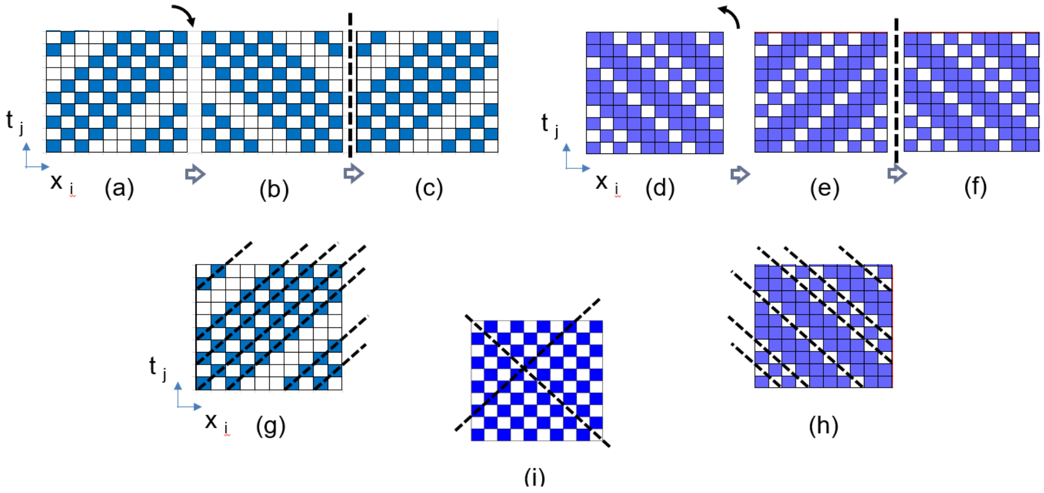


Figure 3: S4 CA-184. The blue cells and the white cells indicate respectively $U_i^j = 1$ and $U_i^j = 0$. t_j and x_i are respectively discretized time and discretized space. (a), (b) and (c) Checkerboard patterns with the cell density $\rho = 0.4$. (d), (e) and (f) $\rho = 0.7$. (g) $\rho = 0.4$. The cell values are invariant along the dashed lines. (h) $\rho = 0.7$. The cell values are invariant along the dashed lines. (i) $\rho = 0.5$. The cell values are invariant along the dashed lines.

In the case of the cell density $\rho = 0.7$ (Fig. 3 (d), (e) and (f)), starting from the initial pattern, the 90 degrees rotation counterclockwise with respect the origin of the $x_i - t_j$ coordinate system, followed by the reflection with respect to the t_j axis at x_{10} results in the initial pattern. The operation is equivalent to the coordinate transformation $(x, t) \to (x', t') = (t, x)$. As in the previous discussions, the simultaneous Burgers equations require the conditions $v_0 = -1/2$ and $v = v(x - 2v_0 t + c)$, or $v = v(x + t + c)$ which is the left-going traveling wave. This result is consistent with the checkerboard pattern in Fig. 3 (h), where the white cells align along the dashed line $t_j = -x_i - c$ so that the motion of the cell is the left-going traveling wave. From Fig. 2, for the cell density $\rho > 0.5$ in which the white cells align along the line $t_j = -x_i - c$, the motion of the cell is left-going traveling wave.

Recalling the discussions in the string theory that phase transition may exist at the domain between the right-moving and the left-moving domain, let us pay attention to the domain of $\rho = 0.5$ which is the domain between the cell density $\rho < 0.5$ and $\rho > 0.5$. The checkerboard pattern for the $\rho = 0.5$ is the mixture of the right-going pattern and the left-going pattern, in which the blue cells align along the dashed line $t_j = x_i + c$ and the white cells align along the dashed line $t_j = -x_i - c$ (Fig. 3 (i)). To be consistent with the previous discussion, define the domain of the right-going traveling wave as $\rho < 0.5$, define the domain of the right-going traveling wave as $\rho > 0.5$, and define the intermediate domain as $\rho = 0.5$.

3.4 Number of states in the periodic CA-184

The number of states is counted with respect to the periodic CA-184 for examining the implication from the string theory that a discontinuous decrease in entropy possibly occurs at $\rho = 0.5$. The state is defined by the configuration of occupation of the 10 cells aligned in the x_i axis at time t_j. Suppose the cells occupied by 1s and 0s are painted blue and white, respectively. The state with the unique checkered pattern with blue and white is counted as one. There are $2^{10} = 1024$ states for the 10 cells aligned. All cells are grouped into 9 groups by the cell density ρ. Each group is then grouped into a few subgroups by the number of clusters in which the blue cells are in contact with each other to form a single mass. The number of states is counted by subgroup.

It should be noted that there is the repulsion force between the cells in CA-184. If there are two occupied cells in touch and if there is an empty space on the right of those cells, the occupied cell on the right moves right to occupy the empty cell. If movable, two cells never stay in contact in the CA-184 (Fig. 4 (a)). Due to the repulsion force, the stable state of CA-184 consists of the maximum number of movable clusters (Eq. (22)) of occupied cells. Here, the immovable cluster in which all cells are occupied is not our concern. For the cell density $\rho = 0.6$, if the initial number of clusters is 4, the number of states in the stable condition is 4 that is the $N_{max\_m\_CL}$, the maximum number of movable clusters (Fig. 4 (b)). If the cell density $\rho = 0.6$, and if the initial number of clusters is either 3 or 1, the number of states in the stable condition is 4 (Fig. 4 (c) and (d)). Regardless of the initial number of clusters, the number of clusters in the stable state becomes $N_{max\_m\_CL}$. This holds not only for $\rho = 0.6$, but also for arbitrary ρ, such as $\rho = 0.4$ (Fig. 4 (e)).

$$\begin{aligned} &N_{max\_m\_CL} = M\, min\{\rho,\, 1-\rho\} \qquad : Maximum\, number\, of\, movable\, cluster \\ &M = 10 \qquad : Total\, number\, of\, cells \end{aligned} \tag{22}$$

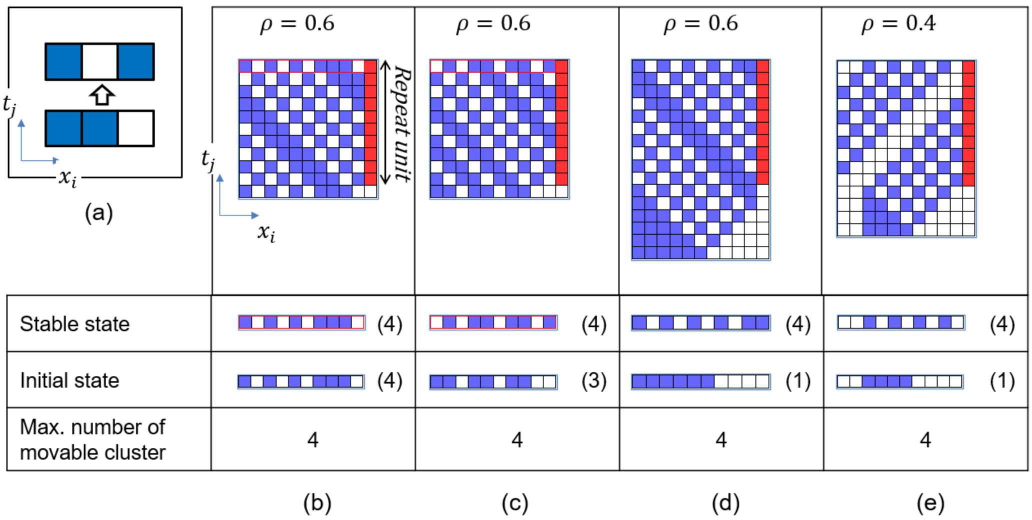
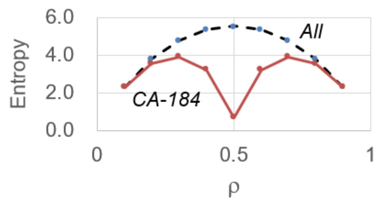


Figure 4: Condensation states in the periodic CA-184. The blue cells and the white cells indicate respectively $U_i^j = 1$ and $U_i^j = 0$. t_j and x_i are respectively discretized time and discretized space. The number in the parentheses in the table indicates the number of clusters. The number of clusters in the stable state is identical to the maximum number of movable clusters. (a) Repulsion between cells. (b), (c) and (d) The checkerboard patterns with the cell density $\rho = 0.6$. (e) The checkerboard pattern with the cell density $\rho = 0.4$.

In order to obtain a general view of the number of states associated with periodic CA-184, all possible initial states are considered, and the number of initial states is counted for each number of clusters $CL_i \leq N_{max\_m\_CL}$ and for each cell density ρ. Then, the sum of all counts indicates the total number of initial states (Eq. (23)), and the number of states for the $CL_i = N_{max\_m\_CL}$ indicates the number of states of the periodic CA-184. Since the number of states for the periodic CA-184 is constantly less than that for the total possible number of states (tables in Fig. 5), it is considered that the states of the periodic CA-184 are condensed due to the repulsive force between cells. The significant condensation of states at $\rho = 0.5$ yields the discontinuous entropy decrease anticipated in the discussions on the string theory (graph in Fig. 5). There is a well-known example that state condensation indicates the phase transition. In the Bose-Einstein Condensate in Helium-4, all states are condensed to the lowest energy state and the viscosity becomes zero below the phase transition temperature (Kittel & Kroemer, 1980). Therefore, we use this analogy to conclude that the periodic CA-184 exhibits a phase transition at $\rho = 0.5$, causing discontinuous state condensation and discontinuous reduction in entropy.

$$
\begin{aligned}
N_t(M\rho) = &\sum_{i=1}^{N_{max\_m\_CL}} N_t(M\rho,\, CL_i) \\
&\rho \quad : Cell\, density\, (Number\, of\, occupied\, cell\, /\, M) \\
&CL_i \le N_{max\_m\_CL} \quad : Number\, of\, cluster \\
&N_t(M\rho,\, CLi) \quad : Number\, of\, states\, for\, given\, \rho\, and\, CL_i \\
&N_t(M\rho) \quad : Number\, of\, states\, for\, given\, \rho
\end{aligned}
\tag{23}
$$



	N.of state per number of clusters and cell density						
CLi →	0	1	2	3	4	5	Subtotal
ρ							
0	1						1
0.1	-	10	-	-	-	-	10
0.2	-	10	35	-	-	-	45
0.3	-	10	60	50	-	-	120
0.4	-	10	75	100	25	-	210
0.5	-	10	80	120	40	2	252
0.6	-	10	75	100	25	-	210
0.7	-	10	60	50	-	-	120
0.8	-	10	35	-	-	-	45
0.9	-	10	-	-	-	-	10
1	-	1					1
Subtotal	1	91	420	420	90	2	1024

N_max_m_CL	N.of states of all states	N.of states for N_max_m_CL
1	10	10
2	45	35
3	120	50
4	210	25
5	252	2
4	210	25
3	120	50
2	45	35
1	10	10

Entropy of all states	Entropy of CA-184
2.3	2.3
3.8	3.6
4.8	3.9
5.3	3.2
5.5	0.7
5.3	3.2
4.8	3.9
3.8	3.6
2.3	2.3

Figure 5: Entropy decrease in the periodic CA-184. Entropy of all states and of periodic CA-184 are plotted with respect to the cell density ρ. Tables show the results of the counting on the number of states.

4 Reproduce the Fourier spectrum immediately before an earthquake

The state condensation in the periodic CA-184 needs to be relaxed for providing a sufficient number of states enough to simulate the complex seismic vibrations. Since the condensation is caused by the repulsive force, or the cell advance rule of the rule-184

(Fig. 4 (a) and Eq. (5)), the repulsive force is determined by the moving probability p of the p-CA-184 (Eq. (8)). We re-define the p as the state condensation factor: $p = 1$ indicates the highest condensation state in which p-CA-184 is identical to CA-184, and $p \leq 1$ indicates the relaxed state of p-CA-184. The phase transition occurs at $(\rho, p) = (0.5, 1.0)$.

4.1 Wave generation with periodic *p*-CA-184

In the periodic p-CA-184, the Fourier transform is applied to the calculated chronological data of the stick-slip timing which is defined as the time when an integer 1 reaches the end of the spatial periodic boundary.

$$
\begin{aligned}
Stick - slip\ sequence\ :\ (t_j,\, N_{ss\_j}) \\
N_{ss\_j} = \begin{cases} 1 & if\ t_j = t_{ss} \\ 0 & if\ t_j \neq t_{ss} \end{cases} \\
t_{ss} = \{t | U^{t-1}_{x=10} = 0,\ U^{t}_{x=10} = 1\} \qquad Stick - slip\ timing
\end{aligned}
\tag{24}
$$

where t_j is discretized time. A fast Fourier transform (FFT) with a block size of 1024 and no filter and overlap is applied to the $(t_j,\, N_{ss\_j})$ data for the periodic p-CA-184 with $(\rho, p) = (0.4, 0.7)$ (Fig. 6 (a) and Fig. 6 (b)). We consider an earthquake as a phase transition at $(\rho, p) = (0.5, 1.0)$. The condition $(\rho, p) = (0.4, 0.7)$ has been selected since we are interested in the pre-earthquake event near (0.5, 1.0) rather than the earthquake itself corresponding to $(\rho, p) = (0.5, 1.0)$. Then, the Stick-slip sequence $(t_j,\, N_{ss\_j})$ (Fig. 6 (a)) is shifted one unit toward the t_j axis and added to the original sequence to generate the linear combination. The FFT result for the linear combination of the periodic p-CA-184 is shown in Fig. 6 (c).

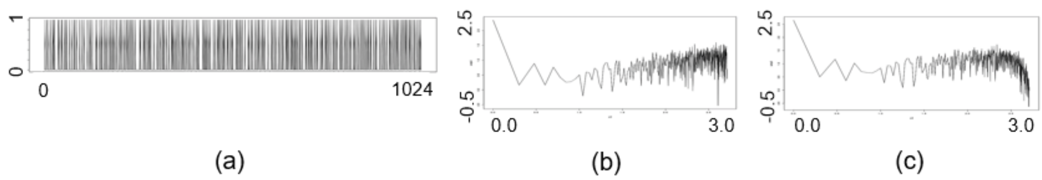


Figure 6: Fast Fourier transform of periodic p-CA-184 $(\rho, p) = (0.4, 0.7)$. (a) Stick-slip sequence $(t_j,\, N_{ss\_j})$. (b) $log_{10} - log_{10}$ plot of Fourier amplitude and frequency for the sequence $(t_j,\, N_{ss\_j})$. $0Hz < frequency \leq 10Hz$. (c) $log_{10} - log_{10}$ plot of Fourier amplitude and frequency for the linear combination of the stick-slip sequence and its shifted sequence. $0Hz < frequency \leq 10Hz$.

On the other hand, the vertical ground vibration velocity data acquired every 0.05 seconds at the seismic station at latitude 38.97 north and longitude 141.53 east is downloaded in chronological order from the web site of National Research Institute for Earth Science and Disaster Resilience (NIED, 2019). The period of the downloaded data is from March 3, 2011 to 11:34 March 9, 2011 that is up to two days before 14:46 on March 11, 2011 when the magnitude 9 Great East Japan Earthquake (GEJE) occurred. The epicenter of the GEJE is at 38.06 north latitude and 142.51 east longitude, approximately 132km away from the seismic station. The downloaded data is divided into blocks of 1024 data corresponding to the acquisition time of 51 seconds and FFT is applied. The upper bound of the frequency domain of the FFT amplitude is 10Hz which corresponds to the half of the data acquisition frequency.

The FFT results for the periodic p-CA-184 (Fig. 6 (c)) is compared to FFT results for the measured seismic data (three FFT plots on the far left in Fig. 7 (c)). The negative curvature at high frequency in the three FFT plots on the far left in Fig. 7 (c) is coincide with the negative curvature at high frequency in the FFT plot in Fig. 6 (c). On the other hand, the weak FFT peaks pointed by the three arrows in Fig. 7 (c) does not exist in Fig. 6 (c). The weak peaks around 0.2Hz are observed not only in the pre-earthquake zone at Kesennuma-shi, where the seismic station is located, but also in other locations of Japan during quiet periods (Fig. 8). The FFT peaks of 0.21Hz to 0.23 Hz observed in all of the three locations and time in Fig. 8, are probably due to the pulsating cycle of the earth caused by ocean waves (Bromirski, P. & Duennebier, F., 2002), and therefore are ignorable.

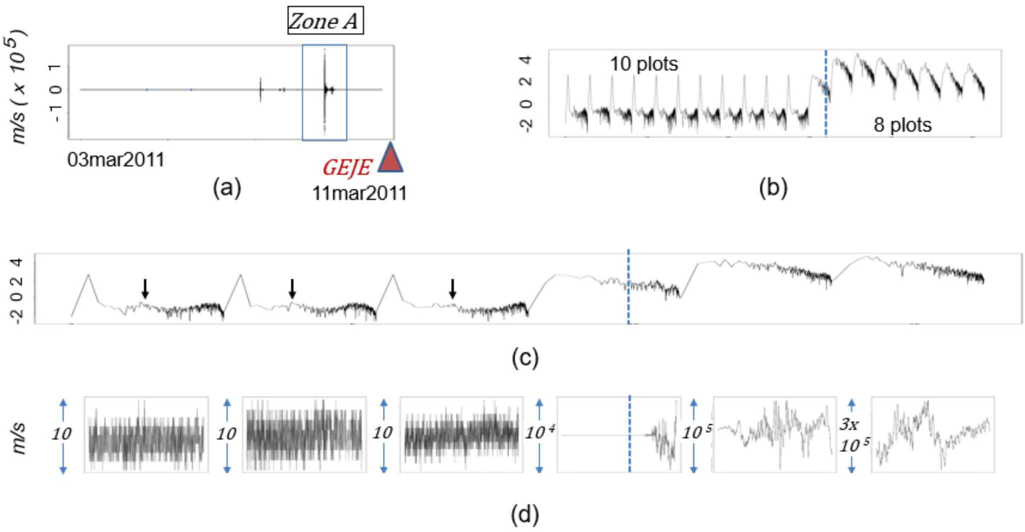


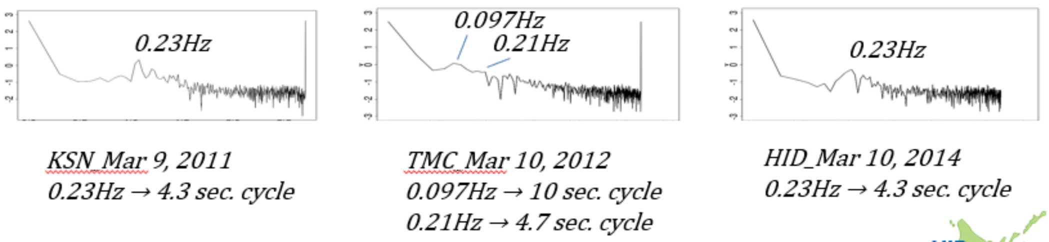
Figure 7: Fast Fourier transform of the seismic data acquired at Kesennuma-shi, Miyagi prefecture, Japan from March 3, 2011 to March 9, 2011(a) Vertical ground vibration (m/s) and time plot. Zone A indicates the period during which an earthquake of magnitude 7.3 was detected. (b) $log_{10} - log_{10}$ plot of the Fourier amplitude and frequency for the data near the earthquake in Zone A. The frequency range of the 18 FFT plots is 0Hz to 10Hz. (c) Enlarged view of the plots in (b). (d) Vibration velocity and time plots corresponding to the plots in (c). The time rage is 0 to 51.2 seconds.

5 Reproduce the Fourier spectrum immediately after an earthquake

The p-BCA (Eq. (9)), which considers the wave shape, is required to reproduce the FFT plot for the post-earthquake since the post-earthquake vibrations of which intensity are 10^4 greater than the pre-earthquake intensity show characteristic wave forms (two plots on the far right in Fig. 7 (d)).

5.1 Wave generation with p-BCA

The wave form which is required as the initial condition in the p-BCA calculation is generated by the difference Burgers equation (Eq. (2)). Solving Eq. (2) near the constant $c = 1$ evolves the given sine waveform into a sharp and asymmetrical shape that



Code	Latitude	Longitude	Prefecture	City
KSN	$38.97N$	$141.53E$	$Miyagi$	$Kesennuma - shi$
HID	$42.82N$	$142.41E$	$Hokkaido$	$Hidaka - cho$
TMC	$32.60N$	$130.91E$	$Kumamoto$	$Shimomashiki - gun$

Figure 8: Earth pulsation. The log_{10} $-$ log_{10} graphs of the Fourier amplitude and frequency for the vibration data acquired at three seismic stations during the quiet period when no significant vibration was detected. The frequency range is 0Hz to 10Hz. KSN, HID and TMC are the code name of the seismic stations, shown in the table with their locations.

resembles a triangle wave. The constant $c < 1$ in Eq. (2) generates a left-moving wave, $c > 1$ generates a right-moving wave, and $c = 1$ generates a standing wave. From the previous discussions on the periodic CA-184 that the phase transition occurs at the domain between the domain of the left-moving wave and the right-moving wave, the $c = 1$ generates the phase transition solution of the difference Burgers equation. We consider the difference Burgers equation with $c = 1$ to correspond to the earthquake and to the periodic CA-184 with $\rho = 0.5$. We generate the initial condition with the difference Burgers equation (Eq. (2)) of $c = 1$ (solid line in Fig. 9 (a)). Then, calculate the evolution of the initial condition due to the probabilistic fluctuation, by using the p-BCA (Eq. (9)) with $\overline{M} = \overline{L} = 133$ and $p = 0.95$. The stable solution, which represent the ground vibration near the earthquake, is shown in Fig. 9 (b).

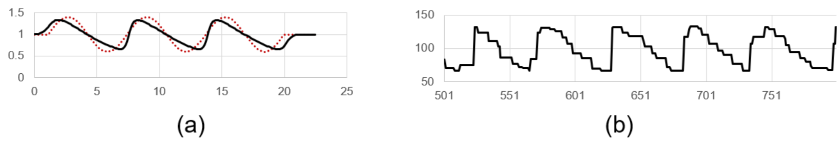


Figure 9: Input and output of p-BCA. (a) Input sine wave (dotted line) and output (solid line) of the difference Burgers equation. The solid line is the initial condition for the p-BCA calculation. (b) Result of the p-BCA calculation.

The Fourier amplitude, with block size 1024, calculated for the result of the p-BCA (Fig. 9 (b)) is compared to the Fourier amplitude of the measured seismic data. The linearly decreasing trend (straight solid line) in the $log_{10} - log_{10}$ plot of the Fourier amplitude and frequency is consistent in both the p-BCA calculation and the measured data (Fig. 10).

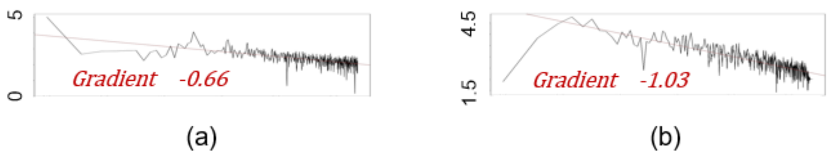


Figure 10: Comparison of p-BCA and measurement. (a) $log_{10} - log_{10}$ plot of the Fourier amplitude and frequency for the result of the p-BCA calculation. The frequency range is 0Hz to 10Hz. (b) $log_{10} - log_{10}$ plot of the Fourier amplitude and frequency for the measured data (extracted from the plot on the far right in Fig. 7 (c)). The frequency range is 0Hz to 10Hz.

6 Gumbel distribution of periodic p-CA-184 near phase transition

We have shown that there exists a phase transition in the periodic CA-184, and that the ground vibration near the earthquake is represented by the p-CA-184 and the p-BCA, which are a family of Burgers equation fluctuating around the phase transition. In particular, the pre-earthquake vibration is reproduced by the p-CA-184. Therefore, an earthquake can be considered as a phase transition associated with the periodic CA-184, and the earthquake precursor can be found in the p-CA-184, if any.

In order to derive earthquake precursors measurable in daily earthquake monitoring, we introduce the Gumbel distribution which is frequently applied to predicting the rare event such as the life time of a steel structure (Yamamoto, M. & Shibata, T., 2002), and investigate the Gumbel distribution in the periodic p-CA184.

The stochastic variable x' in the Gumbel distribution is the time interval D_{ts_j} (Eq. (25)) between the two stick slips evaluated from the stick-slip sequence $(t_j,\, N_{ss\_j})$ (Eq. (24)) calculated for the periodic p-CA-184. The cumulative distribution $G_0(x')$ calculated as shown in Table 1 is fitted to the second equation in Eq. (20) to obtain the gradient, the y-intercept and $R2$, which is the square of the Pearson's correlation coefficient.

$$\begin{aligned} &Stochastic\,variable\,of\,Gumbel\,distribution\,for\,p - CA - 184 \\ &D_{ts\_j} \equiv t_{s\_j} - t_{s\_(j-1)} \qquad Time\,interval\,between\,stick - slips \\ &t_{s\_j} = \{t_j |\, (t_j,\, N_{ss\_j}) = (t_j,\, 1)\} \end{aligned} \tag{25}$$

The earthquake precursors are proposed by evaluating the three regression parameters, gradient, y-intercept and $R2$, for the cell density ρ from 0.1 to 0.9 with the fixed condensation factor of $p = 0.9$.

The $R2$ is maximum at $\rho = 0.4$ and second largest at $\rho = 0.6$, but is relatively small at $\rho = 0.5$, which is the closest to the phase transition (Fig. 11 (b)). Large $R2$ values are expected near the phase transition of the periodic p-CA-184, therefore a large $R2$ is expected before the earthquake (first in Eq. (26)).

The y-intercept and the gradient show large values in $0.4 \leq \rho \leq 0.6$ (Fig. 11 (d) and (f)) if we ignore $\rho = 0.1$ and $\rho = 0.9$, which are far from the phase transition. Therefore, a large absolute value of the product consisting of gradient, y-intercept and $R2$ is expected near the phase transition of the periodic p-CA-184, and is expected before the earthquake. Name the absolute product RGI (second in Eq. (26)).

$$\begin{aligned}
&Earthquake\, precursors:\\
&\qquad 1)\, Large\, value\, of\, R2\, in\, Gumbel\, interpolation\\
&\qquad 2)\, Large\, value\, of\, RGI \equiv |gradient \times y - intercept \times R2|\\
&\qquad\quad in\, Gumbel\, interpolation
\end{aligned} \tag{26}$$

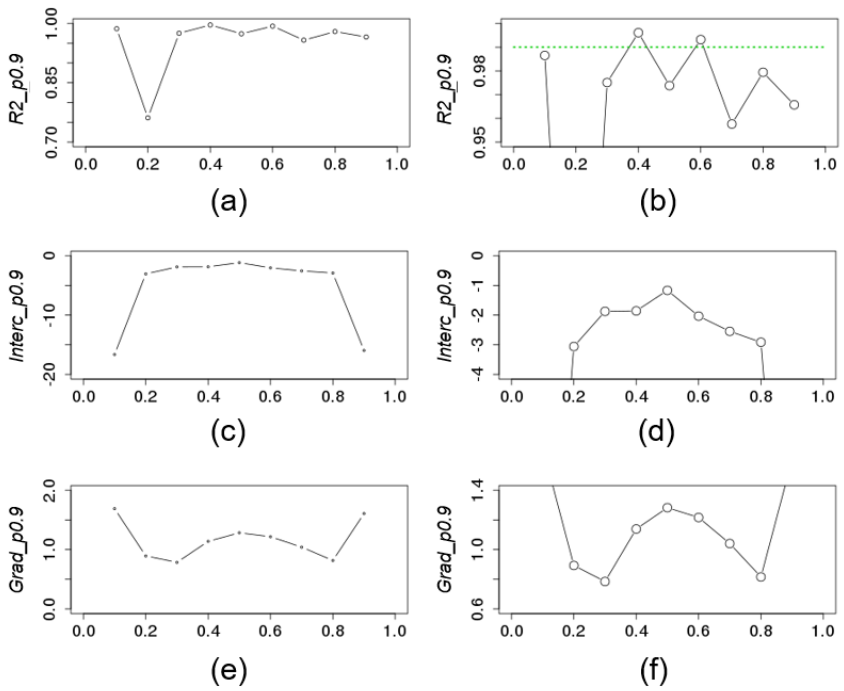


Figure 11: Gumbel interpolation parameters for $p = 0.9$. (a) Plot of (ρ, $R2$). (b) Enlarged view of (a). (c) Plot of (ρ, $y - intercept$). (d) Enlarged view of (c). (e) Plot of (ρ, $gradient$). (f) Enlarged view of (e).

7 Reliability

The reliability of the two precursors of earthquake (Eq. (26)) has been confirmed.

7.1 $R2$ precursor

The $R2$ precursor, the first of the Eq. (26) has been searched to detect earthquakes with seismic intensity greater than 4. The Fourier transform is applied to ground vibration data acquired at HID, TMC and KSN locations for approximately 10 years. The stochastic variable is the gradient in the low frequency zone, from 0.1Hz to 2.9Hz, of the $log_{10} - log_{10}$ plot of the Fourier amplitude with respect to frequency. In the Gumbel interpolation at time t, the entire history of the data from the start of acquisition up to the time t, is considered. In other words, the timescale in the Gumbel interpolation varies from 8.5 hours to 10 years. The calculation details are shown in Appendix A.

The results of searching for the $R2$ precursor are shown in Fig. 12 (a), (b), and (c). In Fig. 12, $minGR2$ is the parameter indicating the large $R2$ with small deviation, which is essentially similar to the $R2$. Therefore, the large $minGR2$ is the precursor of earthquakes. The $P-$ and $D-$ beside the arrows pointing the timing of events in Fig. 12 respectively indicate the earthquake precursor and the detected earthquake. The uppercase alphabets are the index of the earthquakes explained in the tables in Fig. 12 (a'), (b'), and (c'). The tables include the information about earthquakes searched on the website of the Japan Meteorological Agency of Ministry of Land, Infrastructure, Transport and Tourism (JMA, 2019). The search conditions are seismic intensity greater than 4, the location of observing the seismic intensity, and the time period specified in Fig. 12 (a), (b), and (c). The observation points of the seismic intensity are selected near the seismic stations and are described in the caption of the figures.

The $R2$ precursor is detected before every earthquake with the seismic intensity greater than 4 at the seismic stations during the observation period (Fig. 12 (a), (b), and (c)). If the magnitude of those earthquakes is less than 8, the precursor is observed about a month before the earthquake (Fig. 12 (a) and (b)).

In the case of the large earthquake, with a magnitude of 9 at KSN, the precursor is observed 17 months before the earthquake (Fig. 12 (c)). In Fig. 12 (c), the precursor corresponding to the $D-A$ is not shown since the circled zone where the precursor is expected is numerically unstable. In addition, once the $minGR2$ deviates from 1.0 by a large amount after the $D-B, C$, it does not recover to the pre-quake levels. We need to restart the calculation of $minGR2$ after the $D-B, C$. Therefore, the plot after the $D-B, C$ does not reflect the reality, and to avoid misunderstanding, the earthquake or the up arrow is not displayed after the $D-B, C$ in the figure.

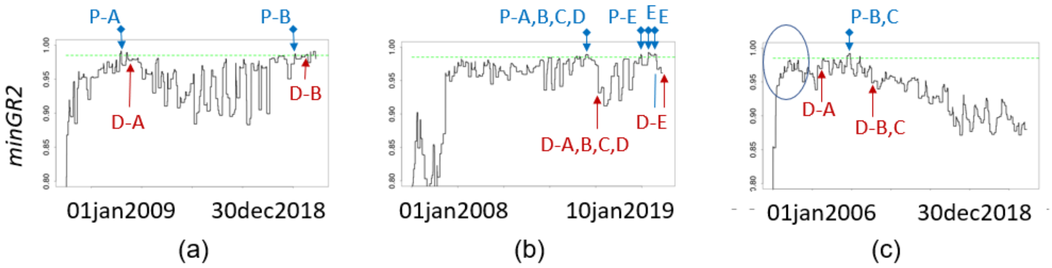
7.2 *RGI* precursor

The RGI precursor, the second of the Eq. (26) has been searched to detect a magnitude 9 earthquake. The Fourier transform is applied to the same ground vibration data as the $R2$ precursor evaluation data. The stochastic variable is the curvature in the high frequency zone, from 2.9Hz to 9.8Hz, of the $log_{10}-log_{10}$ plot of the Fourier amplitude with respect to frequency. In the Gumbel interpolation at time t, the data acquired during approximately 8.5 hours centered on t are used. The calculation details are shown in Appendix B.

In Fig. 13, which is the results of the RGI calculation, the $P-$ and $D-$ beside the arrows respectively indicate the earthquake precursor and the detected earthquake. The uppercase alphabets are the index of the earthquakes explained in the tables in Fig. 12 (a'), (b'), and (c').

RGI values are not significant at HID (Fig. 13 (a)) and TMC (Fig. 13 (b)), in which the maximum magnitude of earthquakes detected over the time period is 7.3 as shown in the tables (a') and (b') in Fig. 12. On the other hand, prominent peak is observed at KSN (Fig. 13 (c)) where the maximum magnitude detected over the time period is 9 (Fig. 12 (c')). The RGI precursor is detected only before the magnitude 9 earthquake.

The $minGR2$ precursor and GRI precursor are respectively observed 17 months and 7 months before the magnitude 9 earthquake indicated by the up arrow of $D-B, C$ in Fig. 13 (c'), so that both are the qualified predictor of the large earthquake. It should be noted that $minGR2$ is calculated from the low frequency data with the varying timescale from 8.5 hours to 10 years, and that GRI is calculated from the high frequency data with the fixed timescale of 8.5 hours. Regardless the significant differences in the calculation condition, both $minGR2$ precursor and GRI precursor are observed at the similar timing before the large earthquake. Therefore, both the $minGR2$ precursor and the GRI precursor are the earthquake precursor and are neither accidental nor noise signals.



(a')

ID	Date	Magnitude	Seismic Intensity	Latitude	Longitude	Depth(km)
A	Sep. 7, 2011	5.1	5	$42.15N$	$142.35E$	10
B	Sep. 6, 2018	6.7	5	$42.41N$	$142.00E$	37

(b')

ID	Date	Magnitude	Seismic Intensity	Latitude	Longitude	Depth(km)
A	Apr. 14, 2016	6.5	5	$32.44N$	$130.48E$	11
B	Apr. 16, 2016	7.3	6	$32.45N$	$130.45E$	12
C	Apr. 16, 2016	5.4	5	$32.45N$	$130.45E$	15
D	Apr. 16, 2016	5.9	5	$32.51N$	$130.53E$	11
E	Jan. 3, 2019	5.1	5	$33.01N$	$130.33E$	10

(c')

ID	Date	Magnitude	Seismic Intensity	Latitude	Longitude	Depth(km)
A	Jul. 24, 2008	6.8	5	$39.43N$	$141.38E$	108
B	Mar. 11, 2011	9	6	$38.06N$	$142.51E$	24
C	Apr. 07, 2011	7.2	5	$38.12N$	$141.55E$	66
D	May 13, 2015	6.8	5	$38.52N$	$142.09E$	46

Figure 12: Reliability of $R2$ precursor. Plot of $minGR2$, which is equivalent to the $R2$, and time measured at the three seismic stations shown in Fig. 8. The down and up arrows respectively indicate the timing of the precursors and the timing of the detected earthquakes. The "$P\ -$ " and "$D\ -$ " in the plots indicate "Precursor-" and "Detected-", respectively. The tables (a'), (b'), and (c') are descriptions of the detected earthquakes in the plots (a), (b), and (c), respectively. (a) Plot for HID. (b) Plot for TMC. (c) Plot for KSN. (a') The seismic intensity is recorded at Sinhidaka-cho, approximately 50km from HID. (b') The seismic intensity is recorded in Kumamoto-kita-ku, approximately 39km from TMC. (c') The seismic intensity is recorded at Kesennuma-shi, approximately 10km from KSN.

8 Conclusions

In deriving the instrumentally measurable earthquake precursors by considering earthquakes as a phase transition phenomenon in CA-184, the followings are concluded.

The phase transition occurs at cell density 0.5 in CA-184, where the number of states or the entropy is abruptly reduced to generate the state condensation.

The ground vibration state near the earthquake is reproduced by p-CA-184 and p-BCA, a family of Burgers equation that fluctuate around the phase transition. In particular, the pre-earthquake vibration is reproduced by the p-CA-184. Consequently, an earthquake is considered as the phase transition of the periodic CA-184, and the earthquake precursor is considered as the precursor of the phase transition of the periodic CA-184.

Large $R2$ and GRI, two precursors of the phase transition of the CA-184, therefore the earthquake precursors, are derived with p-CA-184 by introducing the Gumbel distribution. $R2$ is the correlation coefficient for Gumbel interpolation, and GRI is the absolute value of the product of the Gumbel interpolation parameters $R2$, $Gradient$, and

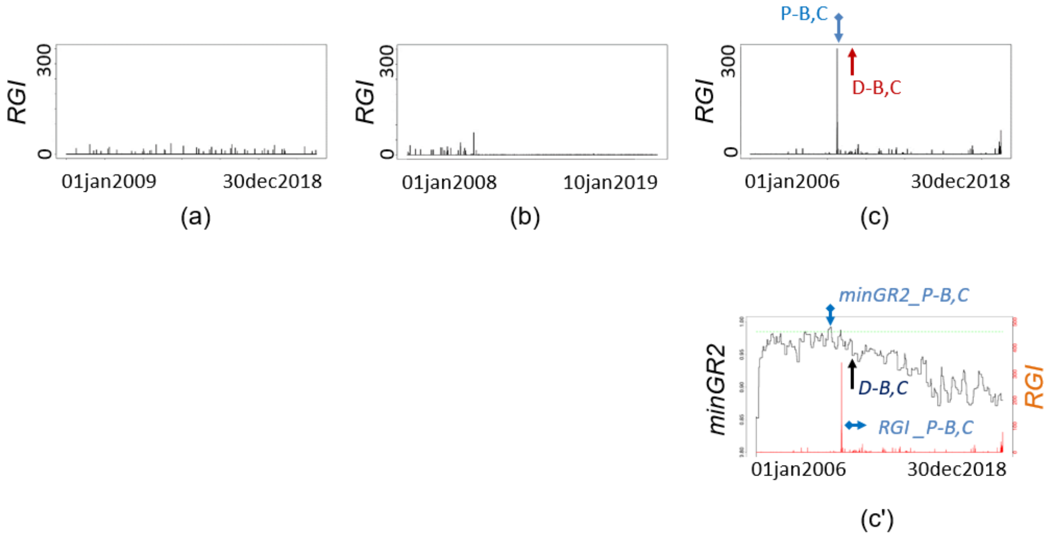


Figure 13: Reliability of RGI precursor. Plot of RGI and time measured at the three seismic stations shown in Fig. 8. The down and up arrows respectively indicate the timing of the precursors and the timing of the detected earthquakes. (a) Plot for HID. (b) Plot for TMC. (c) Plot for KSN. (c') Simultaneous plot of $minGR2$ and RGI with respect to time. Plot for KSN.

$y-intercept$. The $R2$ precursor predicts the earthquake with the seismic intensity greater than 4, while the RGI precursor predicts the large earthquake of magnitude 9.

Two precursors, which are calculated for different frequencies and time scales, are observed at similar timing prior to a magnitude 9 earthquake. The phase transition model of earthquake provides the practical and reliable earthquake prediction method.

Appendix A Details in calculating $R2$

The $R2$ precursor, the first of the Eq. (26) has been searched to detect earthquakes with seismic intensity greater than 4. We have investigated the time evolution of stochastic distributions in the process of earthquakes by introducing the two-step fitting method. The Fourier Transform is applied to the ground vibration data, then the transformed data are fitted to the power-law to extract the power-law parameters, then the parameters are in turn fitted to the Gumbel distribution.

A1 Power-law fitting

We focus on the low frequency domain (Fig. A1). The α, the power-law parameter is obtained by linear regression, and is selected as the stochastic variable to be evaluated in the next step.

A2 Gumbel fitting with extreme value theory

Let us define the time t and the corresponding discrete time t_r such that $t_r \equiv 600r\Delta t_\alpha$ if $600r\Delta t_\alpha < t \leq 600(r+1)\Delta t_\alpha$, where r and Δt_α denote the natural number and the time interval for acquiring the 1024 vibration data for calculating one α, respectively so that the time interval $t_{r+1} - t_r$ is $600\Delta t_\alpha$ corresponding to 5.8 hours in acquiring the

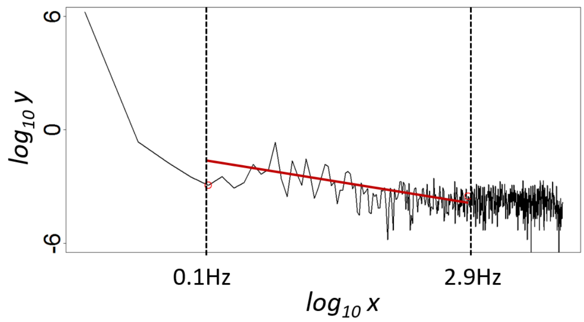


Figure A1: Power-law fitting on the Fourier amplitude calculated from the ground vibration data. The red line illustrates the result of the linear regression analysis in the low frequency domain.

ground vibration data. At the discrete time t_r, the $600r$ sets of α calculated by the power-law fitting for the time interval $0 < t \leq 600r\Delta t_\alpha$ are divided into 20 subsets, each of which contains $30r$ sets of α. Define the M_{30r} as the $30r$ sets of α, and the w as the set of 40 real intervals:

$$\begin{aligned} M_{30r} &= \{\alpha_i |\, 1 \leq i \leq 30r\} \\ \omega &= \{\omega_j |\, [-2.0 + (j-1) \times 0.1, -2.0 + j \times 0.1),\, j = 1, 2, \cdots, 40\} \\ &\qquad \omega_j \equiv [\omega_{j1},\, \omega_{j2}). \end{aligned} \tag{A1}$$

For each ω_j, count the number of $\alpha_i \in M_{30r}$ if the α_i are included in the ω_j, and let the counted result be N_j. From each $30r$ pairs of the (N_j, ω_j), the pair with the maximum ω_j with a non-zero N_j is extracted to form the new set M_{20} consisting of 20 pairs of (N_j, ω_j). It needs to be noted that the size of the $30r$ pairs continuously increase as time goes on, in other words, the time scale in evaluating the Gumbel distribution continuously increases as time goes on.

Each interval, $\omega_j \in M_{20}$, is represented by its mean, $x_j \equiv (\omega_{j1} + \omega_{j2})/2$, which is accompanied by its number of occurrence, N_j. The cumulative distribution, $G_0(x_j)$, is calculated from N_j as shown in Table 1, where N_{sj} denotes the cumulative frequency, $G_0(x_j) \equiv N_{sj}/N_{total}$ is the cumulative distribution, and $N_{total} \equiv \sum_{i=1}^{40} N_i$. The 20 pairs of $(x_j, G_0(x_j))$ are fitted to the Gumbel distribution, and $GR2$, the square of Pearson's correlation coefficient is calculated by linear regression.

Finally, define $minGR2$ as the minimum value in every 100 points in the time series of the $GR2$ data. For the chronological $GR2$ data $\{GR2_1, GR2_2, \cdots, GR2_i, \cdots\}$, the series of the $minGR2$ in chronological order $\{minGR2_1, minGR2_2, \cdots, minGR2_i, \cdots\}$ is defined as follows:

$$minGR2_i \equiv min\{GR2_j \,|\, i - 99 \leq j \leq i\}, \quad i > 99, \tag{A2}$$

so that the first 99 data of the $GR2$ which corresponds to the ground vibration data acquired during the period of 35.06 days are discarded in the $minGR2$ plot. The $minGR2$ is calculated every 5.8 hours at t_r except for the period corresponding to the first 99 data.

The $minGR2$ represents the degree of the correlation in the Gumbel fitting and the fluctuation of the correlation coefficient ($GR2$) from its upper bound, 1.0. A large value in $minGR2$ indicates the high correlation in the Gumbel fitting and the small fluctuation in the correlation coefficient.

Appendix B Details in calculating *RGI*

The RGI precursor, the second of the Eq. (26) has been searched to detect a magnitude 9 earthquake. The curvature of the plot of Fourier amplitude and frequency is introduced to characterize the large earth quakes. The Fourier Transform is applied to the ground vibration data, then the curvatures of the $log_{10}(Fourier\ amplitude){-}log_{10}(Frequency)$ plot are calculated, and then the curvatures are fitted to the Gumbel distribution.

B1 Curvature of the Fourier amplitude

The curvature α is quantitatively defined in Fig. B1. The FFT generates 512 sets of Fourier amplitude from the 1024 sets of the ground vibration data so that each plot in Fig. B1 consists of 512 data points or the 512 pairs of the frequency and the Fourier amplitude. Define the pair $FT_{\_i}$ and the average of the pairs from the j^{th} to the k^{th} of the 512 pairs, $< FT >_{\_(j,k)}$, as:

$$
\begin{aligned}
FT_{\_i} &\equiv (frequency, Fourier\ amplitude)_{\_i} \\
&\equiv (frequency_{\_i}, Fourier\ amplitude_{\_i}) \\
< FT >_{\_(j,k)} &\equiv \left(\sum_{i=j}^{k} \frac{frequency_{\_i}}{(j-k+1)}, \sum_{i=j}^{k} \frac{Fourier\ amplitude_{\_i}}{(j-k+1)} \right).
\end{aligned} \tag{B1}
$$

Calculate $< FT >_{\_(150,154)} \equiv (x_1, y_1) \equiv P_1$ and $< FT >_{\_(500,504)} \equiv (x_2, y_2) \equiv P_2$ so that L_1 and L_2 in Fig. B1 are $x = x_1$ and $x = x_2$, respectively. Here, we focus on the high frequency domain from 2.9Hz ($P1$) to 9.8Hz ($P2$) which consists of 350 data, and the horizontal and the vertical axis of the plot in Fig. B1 are defined as x- and y-axis, respectively. Draw a line L connecting P_1 and P_2, measure the distance L_i between the line L and each of the 350 data points i, extract the maximum value of the L_i, and define it as L_{max_n}. The maximum distance for the plots with the non-negative curvature is similarly defined and denoted as L_{max_p}. The curvature α, which is selected as the stochastic variable to be evaluated in the next step, is defined as:

$$
\begin{aligned}
Curvature\ \alpha \equiv L_{max\_n} &> 0\ for\ the\ negative\ curvature \\
-L_{max\_p} &\leq 0\ for\ the\ non-negative\ curvature
\end{aligned} \tag{B2}
$$

B2 Gumbel fitting with extreme value theory

Let us define the time t and the corresponding discrete time t_r such that $t_r \equiv 600r\Delta t_\alpha$ if $600r\Delta t_\alpha < t \leq 600(r{+}1)\Delta t_\alpha$, where r and Δt_α denote the natural number and the time interval for acquiring the 1024 vibration data for calculating one α, respectively. The time interval $t_{r+1}{-}t_r$ is $600\Delta t_\alpha$ corresponding to 5.8 hours for acquiring the 1024 sets of the ground vibration data. At the discrete time t_r, the 600 sets of α calculated by the power-law fitting for the time interval $600r\Delta t_\alpha < t \leq 600(r+1)\Delta t_\alpha$ are divided into 20 subsets, each of which contains 30 sets of α. Define the M_{30} as the 30 sets of α, and the w as the set of 40 real intervals:

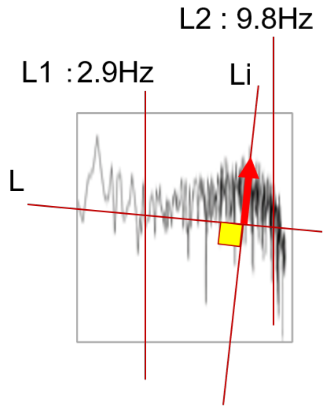


Figure B1: Definition of the curvature α. $\alpha \equiv max\{L_i : 152 \leq i \leq 502\} \equiv L_{max\_n} > 0$. The data between L_1 (2.9Hz) and L_2 (9.8Hz) are evaluated. The red arrow is the L_i. .

$$
\begin{aligned}
M_{30} &= \{\alpha_i | \, 1 \leq i \leq 30\} \\
\omega &= \{\omega_j | \, [-4.0 + (j-1) \times 0.2, -4.0 + j \times 0.2), \, j = 1, 2, \cdots, 40\} \\
&\qquad \omega_j \equiv [\omega_{j1}, \, \omega_{j2}).
\end{aligned} \tag{B3}
$$

For each ω_j, count the number of $\alpha_i \in M_{30}$ if the α_i are included in the ω_j, and let the counted result be N_j. From each 30 pairs of the (N_j, ω_j), the pair with the maximum ω_j with a non-zero N_j is extracted to form the new set M_{20} consisting of 20 pairs of (N_j, ω_j).

Each interval, $\omega_j \in M_{20}$, is represented by its mean, $x_j \equiv (\omega_{j1} + \omega_{j2})/2$, which is accompanied by its number of occurrence, N_j. The cumulative distribution, $G_0(x_j)$, is calculated from N_j as shown in Table 1, where N_{sj} denotes the cumulative frequency, $G_0(x_j) \equiv N_{sj}/(N_{total} + 1)$ is the cumulative distribution, and $N_{total} \equiv \sum_{i=1}^{40} N_i$. The 20 pairs of $(x_j, G_0(x_j))$ are fitted to the Gumbel distribution. Then, the absolute product RGI of the gradient $\frac{1}{\eta}$, the y-axis intercept $\left(-\frac{\lambda}{\eta}\right)$, and the square of Pearson's correlation coefficient $R2$ obtained in the Gumbel-fitting is calculated.

Acknowledgments

Data is publicly available through National Research Institute for Earth Science and Disaster Resilience, National Research and Development Corporation under Ministry of Education, Culture, Sports, Science and Technology. F-Net (Broadband seismograph network) data base. http://www.fnet.bosai.go.jp/top.php?LANG=en.

References

Geller, R. J. (1997). Earthquake prediction: a critical review. Geophysical Journal In-ternational,131,425-450. https://doi.org/10.1111/j.1365-246X.1997.tb06588.x

Nishinari, K. & Takahashi, D. (1999). Analytical properties of ultradiscrete Burgers equation and rule-184 cellular automaton. Journal of Physics A: Mathematical and General. 31. 5439. DOI: 10.1088/0305-4470/31/24/006

Polchinski, J. (1998). String Theory Volume I: Cambridge University Press. (Polchinski, J. Ito, K., Odake, S. & Matsuo, Y.(Translation) (2007). String Theory Volume I: Springer Japan, pp.11-15, pp.39-42, pp.57-58, pp.63-69)

Carlip, S. (1999). Black hole entropy from conformal field theory in any dimension. Phys.Rev.Lett. 82, 2828-2831. DOI: 10.1103/PhysRevLett.82.2828

Carlip, S. (2000). Logarithmic corrections to black hole entropy from the Cardy formula. Class. Quant. Grav. 17, 4175-4186. e-Print: gr-qc/0005017 [gr-qc]. DOI: 10.1088/0264-9381/17/20/

Charras-Garrido, M., Lezaud, P. (2013). Extreme Value Analysis : an Introduction. Journal de la Société Française de Statistique, 154, 2, 66-97. https://hal-enac.archives-ouvertes.fr/hal-00917995

Egghe, L. & Leydesdorff, L. (2009). The relation between Pearson's correlation coefficient r and Salton's cosine measure. Technology, 60 (5), 1027-1036.

Kittel, C. & Kroemer, H. (1980). Thermal Physics: W. H. Freeman and Company San Francisco, pp.199-217.

NIED. (2019). National Research Institute for Earth Science and Disaster Resilience, National Research and Development Corporation under Ministry of Education, Culture, Sports, Science and Technology. F-Net (Broadband seismograph network) data base. http://www.fnet.bosai.go.jp/top.php?LANG=en

Bromirski, P. & Duennebier, F. (2002). The near-coastal microseism spectrum: Spatial and temporal wave climate relationships. Journal of Geophysical Research. 107. 10.1029/2001JB000265.

Yamamoto, M. & Shibata, T. (2002). Program for Calculating the Extreme Value Statistics on the Spreadsheets Software (EVAN-II). Corrosion Science and Technology, Vol. 31, No. 3, pp, 228-233

JMA. (2019). Japan Meteorological Agency under Ministry of Land, Infrastructure, Transport and Tourism. Seismic intensity database search. https://www.data.jma.go.jp/svd/eqdb/data/shindo/index.php