EXPRESS LETTER

Continuous monitoring of instrumental clock errors at 50 volcanoes in Japan based on seismic interferometry

Takashi Hirose[*] and Hideki Ueda

Abstract

The correct absolute time of a seismogram is an important prerequisite for many seismological processing methods. Recently, seismic interferometry has been applied to continuously estimate instrumental clock errors in certain regions. However, previous studies have focused on estimating clock errors in only one target region for each study, and data processing methods that provide more stable results in different regions have not been explored. This study presents a new method to estimate clock errors that are more stable in various regions based on seismic interferometry with windowed cross-correlation and least absolute deviation regression. The applicability of this method was validated by its application to 50 active volcanoes in Japan. We found clock errors at 23 stations at 13 volcanoes between January 2017 and December 2021. The maximum amplitude of the clock errors ranged between 0.24 and 18.27 s. This value range is significantly larger than the amplitude of the background fluctuation in the estimated values of clock errors, which is less than 0.05 s for most volcanoes. A comparison of different approaches for the estimation of clock error shows that our proposed method that uses windowed cross-correlation and least absolute deviation regression can reduce the effect of waveform changes in seismic ambient noise cross-correlation functions, probably because of noise source fluctuations. Managing the time stamps of data using the method proposed in this study will contribute to data quality assurance and ensure the reliability of analysis using time information in various study regions.

Keywords Instrumental clock error, Seismic interferometry, Active volcano
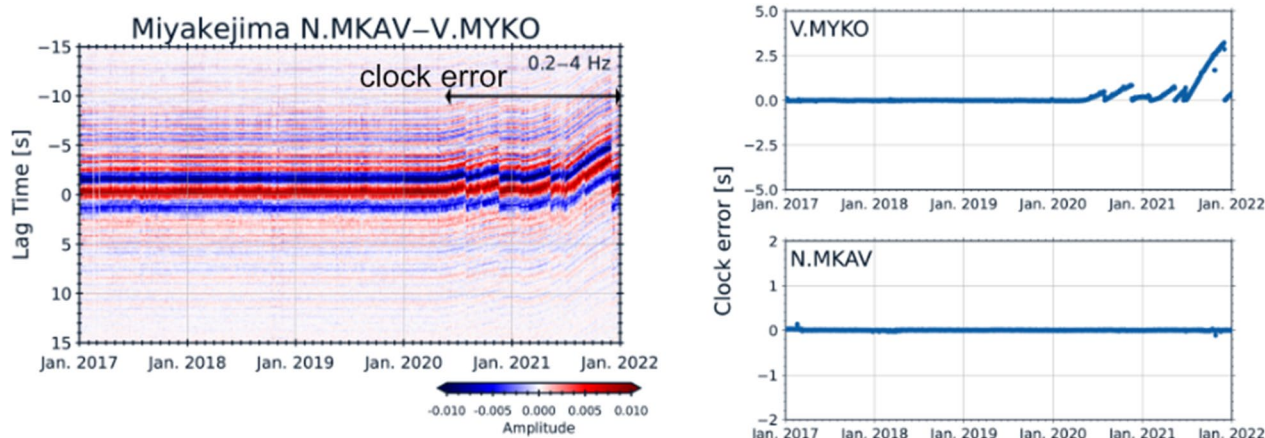
*Correspondence:
Takashi Hirose
takashi.hirose@bosai.go.jp
Full list of author information is available at the end of the article

Graphical Abstract

Introduction

The correction of the absolute time of a seismogram is an important prerequisite for many seismological processing methods. The internal clock of the seismic station is usually synchronized frequently with the global positioning system (GPS) satellite, which serves as a highly accurate external reference clock and ensures precise time. When GPS signals are lost, synchronization may fail, and the internal clock usually begins to drift primarily in relation to the reference time. The evaluation and correction of such clock errors are long-standing problems in seismology.

Recent advances in seismic interferometry (Campillo and Paul 2003; Curtis et al. 2006) have enabled the study of temporal changes in subsurface structures (Brenguier et al. 2014; Wang et al. 2017; Hirose et al. 2020) and background velocity structures. Stehly et al. (2007) demonstrated that seismic interferometry can be useful for the estimation of instrumental clock errors. This approach has been applied to continuous seismograms at ocean-bottom seismometers (OBSs) and seismic stations on land in various regions (Sens-Schonfelder 2008; Gouédard et al. 2014; Takeo et al. 2014; Le et al. 2017; Hable et al. 2018). However, previous studies have not estimated clock errors in different regions using the same data processing method. Because clock error estimation based on seismic interferometry can be widely applied to seismic networks in various regions, such a comprehensive approach and validation will contribute to accelerating applications.

This study estimated instrumental clock errors at 50 active volcanoes in Japan based on seismic interferometry to evaluate the applicability of clock error estimation.

Moreover, we explored suitable data processing methods to stably estimate clock errors at these volcanoes.

Methods

Calculation of seismic ambient noise cross-correlation functions

We first applied the seismic interferometry technique to continuous seismograms at 50 volcanoes and calculated seismic ambient noise cross-correlation functions (CCFs) representing Green's functions between two seismic stations. We used continuous seismic data from the V-net (National Research Institute for Earth Science and Disaster Resilience (NIED)) and Japan Meteorological Agency (JMA) seismic stations for the 50 volcanoes (Fig. 1a). Although more than three seismic stations are necessary to estimate the clock error for each station, only one or two stations are installed at some volcanoes. Therefore, we also used neighboring Hi-net and F-net seismic stations (NIED) around volcanoes when the number of seismic stations was insufficient. Figure 1b shows topographic maps and spatial distributions of seismic stations in the Nasu, Asama, Miyakejima, Kusatsu-Shirane, Kozushima, Hokkaido-Komagatake volcanoes. Clock errors were detected at these volcanoes during the target period, between January 2017 and December 2021 (topographic maps at all 50 volcanoes are displayed on Additional file 1: Figs. S1–S6). The black/red reversed triangles represent the seismic stations. A three-component short-period seismometer (natural frequency of 1 Hz) and/or broadband seismometer is installed at each station. Seismograms were recorded at a sampling frequency of 100 Hz or 200 Hz. Data with a sampling frequency of 200 Hz was downsampled to 100 Hz. We first
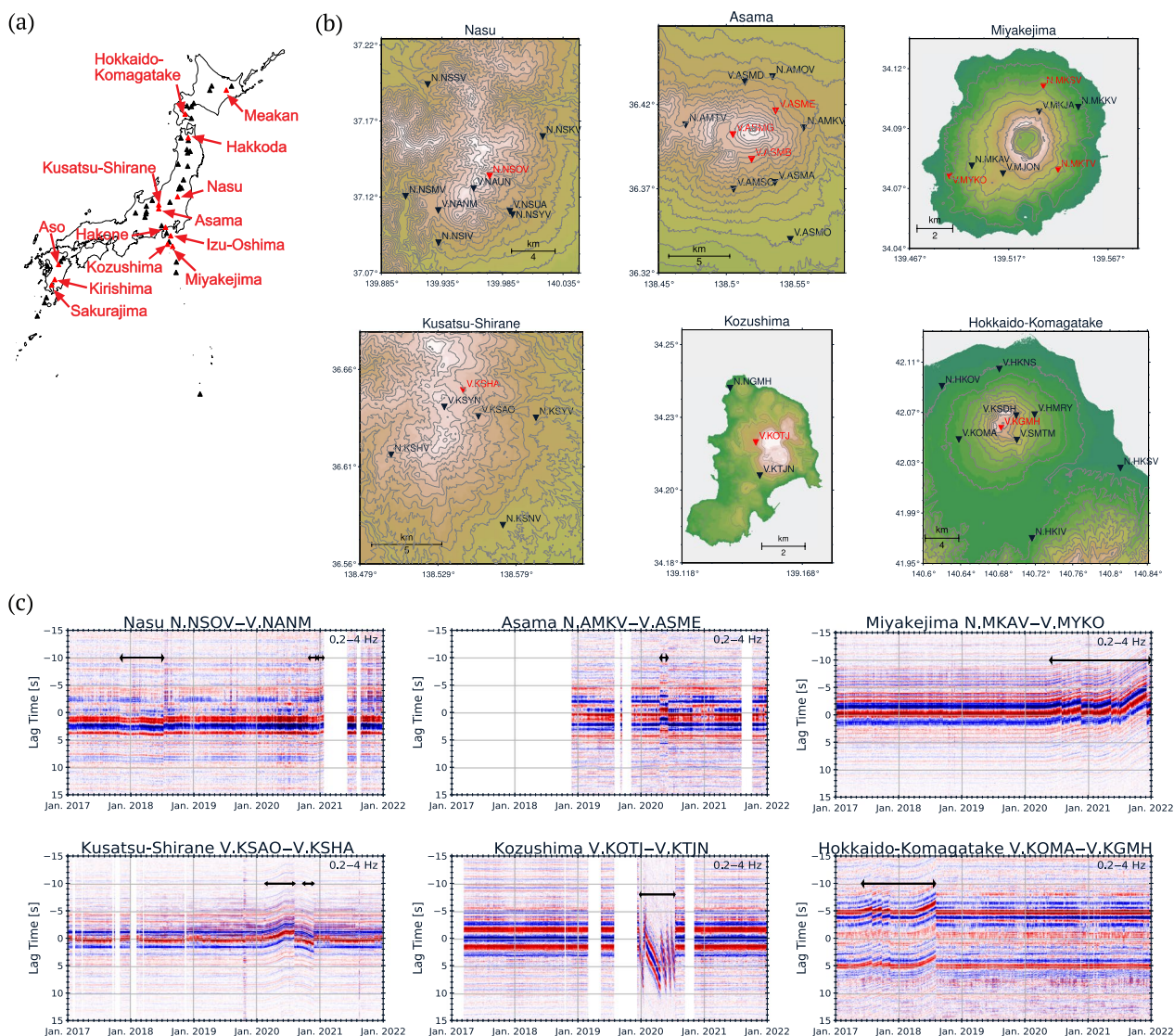
Fig. 1 Topographic maps and examples of stacked cross-correlation functions (SCCF) waveforms. **a** Locations of 50 volcanoes. 13 volcanoes where clock errors occurred during the study period are indicated by red triangles. **b** Topographic maps of 6 volcanoes where clock errors occurred during the study period. The red inverted triangles on the topographic maps represent seismic stations for which clock errors were detected. Note that topographic maps of all 50 volcanoes are shown in Additional file 1: Figs. S1–S6. **c** Examples of temporal changes in SCCF (3-day-stacked CCF) waveforms for station pairs with clock errors. Horizontal black arrows represent the time periods when clock errors occurred (see also Additional file 1: Figs. S7–S9 for other station pairs for which clock errors were detected)

corrected the instrumental response (Maeda et al. 2011) and then applied seismic interferometry to continuous seismograms on the vertical component between January 2017 and December 2021.

We divided the continuous seismic records into 10-min segments and applied spectral whitening to reduce the influence of persistent monochromatic noise sources (Bensen et al. 2007). To improve the temporal stability of noise records before correlation, we also applied one-bit normalization (Bensen et al. 2007), which removes

irregular events. After this preprocessing, we computed the CCFs every 10 min and stacked them over 3 days (SCCFs). Figure 1c shows examples of the temporal changes in the SCCFs. A fourth-order Butterworth filter between 0.2 and 4 Hz was applied to each SCCF. Drifts in SCCFs can be seen for station pairs N.NSOV–V.NANM at Nasu N.MKAV–V.MYKO at Miyakejima, V.KSAO–V.KSHA at Kusatsu-Shirane, V.KOTJ–V.KTJN at Kozushima, and V.KOMA–V.KGMH at Hokkaido-Komagatake

(see horizontal black arrows in Fig. 1c). The SCCFs for the station pair N.AMKV–V.ASME at Asama suddenly increased in April 2020 and recovered in early June 2020. These drifts or jump imply significant clock errors at these seismic stations (see also Additional file 1: Figs. S7–S9 for other examples). Although not discussed in this paper, improving the time resolution of the clock error estimation could be achieved by using the Welch's method in calculating seismic ambient noise CCFs (Seats et al. 2012) and/or applying the denoising method of seismic ambient noise CCFs (e.g., Moreau et al. 2017; Viens and Van Houtte 2020).

Estimation of relative clock errors for each station pair

Gouédard et al. (2014) applied a doublet method (Poupinet et al. 1984) to estimate clock errors in seismograms of OBSs. In the doublet method, the delay times between the seismic ambient noise CCFs are measured by calculating the cross-spectra of these signals. After measuring the delay times for each lag time by sliding a short-time window, a straight line was fitted to the estimated delay times for each lag time. The intercept of the fitted straight line represents the relative clock error between two stations, which causes the same delay times for all delay times between two seismic ambient noise CCFs (Stehly et al. 2007; Sens-Schonfelder 2008). We developed an approach proposed by Gouédard et al. (2014). We first measured the delay times between SCCFs on two different days in the time domain by sliding a short time window along the lag times (hereafter called the windowed cross-correlation; (WCC) method). The WCC parameters (length of the short time window, sliding step in time, and lag time range) are summarized on the "parameters" sheet in Additional file 1: Table S1. We then applied the least absolute deviation (LAD) regression for line fitting and estimated the intercept of the fitted straight line corresponding to the relative clock error between the two stations. This step is the process devised in this study. LAD is a linear regression technique that minimizes the ℓ^1 norm of residuals between observations and the model. This technique is more robust to outliers than ordinary linear least-squares (OLS) regression, which minimizes the sum of the squared residuals. As LAD regression is an optimization problem that includes a nondifferential function, we cannot simply solve it, such as with OLS regression. We used the alternating direction method of multipliers (ADMM) (Boyd et al. 2011) to perform the LAD regression. ADMM is an algorithm that solves convex optimization problems and is a variant of the augmented Lagrangian scheme

that uses partial updates for the dual variables. This algorithm is applied to solve the following problem:

$$\underset{x,z}{\text{minimize}}\, f(x) + g(z),\ \text{subject to}\, Ax + Bz = C. \quad (1)$$

In case of the LAD regression, the problem can be written as

$$\begin{aligned} &\textit{minimize} \quad \| z \|_1 \\ &\text{subject to} \quad \Phi x - z = y. \end{aligned} \quad (2)$$

Here, y is a vector of estimated delay times between two seismic ambient noise CCFs for each time window. x is a vector of model parameters (slope and intercept of fitted straight line). Φ is the Vandermonde matrix. $f(x) = 0$ and $g(z) = \|z\|_1$ (ℓ^1 norm), $A = \Phi$, $B = -1$, and $C = y$, respectively. The corresponding Augmented Lagrangian form (cost function) for the LAD regression can be written as,

$$\begin{aligned} L_\rho(x,z,y) &= f(x) + g(z) + h^{\mathrm{T}}(Ax + Bz - C) + \tfrac{\rho}{2} \| Ax + Bz - C\|_2^2 \\ &= \| z \|_1 + \tfrac{\rho}{2} \| \Phi x - z - y + u\|_2^2 \left(u = \tfrac{h}{\rho}\right). \end{aligned} \quad (3)$$

Here, $\rho (> 0)$ is called the penalty parameter. In this study, we fixed $\rho = 1$ (Boyd et al. 2011). h is called the Lagrange multiplier. That cost function is minimized by updating x, z, and u as follows (Boyd et al. 2011):

$$\begin{aligned} x^{k+1} &:= \left(\Phi^{\mathrm{T}}\Phi\right)^{-1}\Phi^{\mathrm{T}}\left(y + z^k - u^k\right), \\ z^{k+1} &:= S_{\frac{1}{\rho}}\left(\Phi x^{k+1} - y + u^k\right), \\ u^{k+1} &:= u^k + \Phi x^{k+1} - z^{k+1} - y. \end{aligned} \quad (4)$$

Here, $S_{1/\rho}$ is a proximal operator defined as

$$S_{1/\rho}(x) = \begin{cases} x - 1/\rho & (x > 1/\rho) \\ 0 & (-1/\rho \le x \le 1/\rho) \\ x + 1/\rho & (x < -1/\rho) \end{cases}. \quad (5)$$

This type of proximal operator is called a "soft-thresholding function". In this study, we set the initial values of x, z, and u as $x^0 = 0$ (i.e., intercept and slope of fitted straight line equal to 0), $z^0 = 0$, and $u^0 = 0$.

In the estimation of relative clock errors, there is a need to select an SCCF when no clock error occurs as a reference SCCF. In many cases, however, we do not have information in advance as to whether the clock error does not occur on that day. Therefore, we use an approach similar to that used by Brenguier et al. (2014) to estimate temporal changes in relative seismic velocities. They calculated the velocity difference between every pair of CCFs for different days and then inverted these measurements for the best-fitting relative velocity change time-series. Similarly, daily relative clock errors for each

station pair are estimated by conducting a linear least-squares inversion using relative clock errors calculated from all possible combinations of SCCFs:

$$\underbrace{\mathbf{d}}_{K\times 1} = \underbrace{\mathbf{G}}_{K\times L}\underbrace{\mathbf{m}}_{L\times 1}, \tag{6}$$

$$\underbrace{\begin{pmatrix} \delta t\left(SCCF_{\text{day 1,day2}}\right) \\ \delta t\left(SCCF_{\text{day1,day3}}\right) \\ \vdots \\ \delta t\left(SCCF_{\text{day }i,\text{day}j}\right) \\ \vdots \\ \delta t\left(SCCF_{\text{day }L-1,\text{day}L}\right) \end{pmatrix}}_{K\times 1}$$

$$= \underbrace{\begin{pmatrix} 1 & -1 & 0 & \dots & 0 & 0 \\ 1 & 0 & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & -1 \end{pmatrix}}_{K\times 1} \underbrace{\begin{pmatrix} \delta t_{\text{day1}} \\ \delta t_{\text{day2}} \\ \vdots \\ \delta t_{\text{day}i} \\ \vdots \\ \delta t_{\text{day}L} \end{pmatrix}}_{L\times 1} \tag{7}$$

Here, L represents the total number of target days (1826 days). K represents the number of all possible combinations of SCCFs $\left(_{1826}C_2 = 1,666,225\right)$. $\delta t\left(SCCF_{dayi,dayj}\right)$ represents the relative clock error between days i and j. δt_{dayi} on the right-hand side of Eq. 7 represents the relative clock error on the day i. The inverse matrix approach to estimate model parameter vector $\mathbf{m}$ as used in Brenguier et al. (2014) is hard to apply when the data set is large due to the high computation cost. Therefore, we conduct the LAD regression using ADMM when estimating the relative clock error for each station pair.

Estimation of clock errors for each station

If more than three seismic stations are usable, one can estimate the clock errors for each station by conducting linear least-squares inversion with relative clock errors:

$$\underbrace{\begin{pmatrix} \delta t_{12} \\ \delta t_{13} \\ \vdots \\ \delta t_{ij} \\ \vdots \\ \delta t_{(N-1)N} \end{pmatrix}}_{M\times 1} = \underbrace{\begin{pmatrix} 1 & -1 & 0 & \dots & 0 & 0 \\ 1 & 0 & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & -1 \end{pmatrix}}_{M\times N} \underbrace{\begin{pmatrix} \Delta_1 \\ \Delta_2 \\ \vdots \\ \Delta_i \\ \vdots \\ \Delta_N \end{pmatrix}}_{N\times 1} \tag{8}$$

Here, M and N are the numbers of pairs of stations and seismic stations, respectively. δt_{ij} represents the relative

clock error for the pair of stations i and j, and Δ_i on the right-hand side of Eq. 8 represents the clock error at station i. We define a positive (negative) value of Δ_i to be the clock running behind (ahead). In this study, we calculated the signal-to-noise ratios (SNRs) of SCCFs and discarded SCCFs whose SNRs were low before conducting inversions. The SNRs were calculated for the causal and acausal parts of the SCCFs individually and then averaged. We defined the noise level of an SCCF as the root-mean-squared amplitude of the SCCF between 80 and 90 s in lag time (-90 and -80 s for the acausal part). The signal amplitude was defined as the maximum amplitude of the SCCF in the lag time range (see "parameters" sheet in Additional file 1: Table S1). The SNRs are a fraction of these values. In this study, we set the threshold for SNRs to three for Chokaisan and five for the other volcanoes. These thresholds were determined by considering the number of days in which the clock error could no longer be estimated owing to discarding SCCFs. After the data selection, we estimated the clock errors for each seismic station using the LAD regression. Note that the absolute clock error remains undetermined, because a constant offset can be added to all stations simultaneously without changing time differences. This is achieved by modifying the time of the network (i.e., shifting all clock offsets by the same amount) to set the clock offset of the seismic station, which has reliable timings to zero. In this study, we visually checked the inversion result of clock errors and selected some stations whose estimated clock errors were stable in time as reference stations (see also the "parameters" sheet in Additional file 1: Table S1). We assumed the mean of the estimated clock errors for these stations as an offset, and finally subtracted this offset value from the estimated clock errors for each station.

Results

First, we estimated the relative clock errors between the two seismic stations for all available pairs at each volcano. Figure 2a, and b shows an example of the application of the WCC method for the station pair N.AMKV–V.ASME in the Asama volcano. The top panel of Fig. 2a shows the waveforms of the SCCF on the day without a clock error (blue line) and that on the day the clock error occurred (orange line). A clock error causes the same delay times between two SCCFs for all lag times, and the two SCCFs are shifted by approximately 1 s. The bottom panel of Fig. 2a shows the delay times for each lag measured using the WCC method. The solid red line in this panel represents the straight line fitted by the LAD regression. The slope of this straight line was -0.0018, representing almost the same time shifts that occurred throughout the lag times, and the intercept was estimated to be ~ 1.0 s. The straight line from the OLS regression (dashed red

line) is also shown in the same panel. This fitted straight line is obviously affected by the outlier at -18 s in lag time, and the estimated intercept is 0.15 s smaller than that from the LAD regression. Figure 2b is similar to Fig. 2a but for SCCFs without the clock error. In this case, the relative clock error was estimated to be less than 0.02 s.

We estimated the clock errors for each seismic station by conducting a linear least-squares inversion using the relative clock errors for each station pair. We first considered the case in which an amplitude of estimated clock error exceeded 0.05 s for five consecutive days as a candidate of true clock error and then visually checked if the CCF waveform drifts or jumps. The amplitude of background fluctuations was smaller than 0.05 s for most of the volcanoes (see also Tables 1 and 2); therefore, we set the threshold at 0.05 s. Figure 2c shows the temporal changes in the clock error for each station. All stations in which clock errors occurred during the target period are displayed. For example, the N.NSOV clock at the Nasu volcano, maintained by the NIED, started to drift at approximately 0.008 s/day from late December 2020 to January 21, 2021. The clock error reached 0.24 s on January 21, 2021. A malfunction of the GPS antenna that occurred on January 8, 2021 should be related to this drift. The drifts of the N.NSOV clock were also estimated between October 2017 and July 2018. Significant clock errors were also detected in other 12 volcanoes: Asama, Miyakejima, Kusatsu-Shirane, Kozushima, Hokkaido-Komagatake, Hakone, Meakan, Izu-Oshima, Kirishima, Sakurajima, Hakkoda, and Aso. For example, clock errors that reached around 18 s at maximum occurred at stations N.MKSV and N.MKTV in Miyakejima and N.GJKV, N.OOHV, N.OSMV, and N.ODKV in Izu-Oshima during October 2021. These clock errors were caused by data logger malfunctions. The results of the clock error estimation for all 50 volcanoes are summarized in Tables 1 and 2. Furthermore, we provide all values of clock errors for 50 volcanoes (Additional file 1: Table S1).

Uncertainties for estimating clock errors were evaluated using wild bootstrap (Wu 1986) and are shown as error bars in Fig. 2c. The wild bootstrap procedure is as follows: (1) calculate a residual e_i between the observation and model using the estimated model parameter vector $\widehat{\boldsymbol{m}}$, $e_i = d_i - G_{ij}\widehat{m}_j$ ($i=1,2,...,$n). (2) Compute the bootstrap sample by randomly weighting the residuals $d_i^* = G_{ij}\widehat{m}_j + e_i V_i^*$. We used the standard normal distribution, V_i^*. (3) Estimate $\widehat{\boldsymbol{m^*}}$ by the linear least-squares inversion with data vector $\boldsymbol{d}^*$ from the bootstrap sample. (4) Remake V_i^* and repeat (2) and (3) 1000 times. (5) Calculate the standard deviation of $\widehat{\boldsymbol{m^*}}$. Most of the uncertainties in estimating the clock errors at the seismic stations shown in Fig. 2c were less than 0.03 s.

To confirm the validity of the estimated values of the clock errors, we recalculated the seismic ambient noise CCFs after correcting the timestamps of the seismograms using estimated clock errors and then estimated clock errors again. Red circles in Fig. 2c represent estimated clock errors after correcting for the timestamps of the seismograms. We see that the amplitudes of the clock errors after correction are significantly reduced. This indicates that the estimated values of the clock errors are reliable.

Discussion

### Indicator to judge if a clock error has occurred

Evaluating the range of values of the clock error for days without clock errors (hereafter called background fluctuation) will be useful for judging whether the estimated value of the clock error is normal. We calculated 98 percentiles of absolute values of clock errors for each volcano, and they range between 0.01 and 0.07 s (Tables 1 and 2). Hable et al. (2018) estimated the clock errors in Réunion Island based on seismic interferometry, and these time series show almost annual seasonal variations with an amplitude of approximately $\pm$ 0.1 s. Gouédard et al. (2014) estimated temporal changes of clock errors over 1 year for OBS arrays in the Mid-Atlantic-Ridge and the East Pacific Rise, and these time series show temporal variations with an amplitude of approximately $\pm$ 0.04–0.08 s. The amplitude of the temporal variations in clock errors for the 50 volcanoes is of the same order or slightly smaller than those of previous studies.

As mentioned in the Method part, a relative clock error for each station pair corresponds to an intercept of a straight line fitted to estimated delay times between two SCCFs for each lag time. When a significant clock error occurs, fitting a straight line with intercept (i.e., $y = ax + b, b \neq 0$) is more reasonable than fitting a straight line through the origin (i.e., $y = ax$). The Akaike information criterion (AIC) has been widely used as one of the indicators of model selection. We compare AIC for the two models, straight lines with/without intercept. Hurvich and Tsai (1990) proposed a corrected version of AIC for the LAD regression:

$$cAIC = n\left(\log\widehat{\sigma}^2 + 1\right) + n + 2(p+1) + \frac{2(p+1)(p+2)}{n-p-2},$$
$$\widehat{\sigma} = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - x_i'\widehat{\theta}\right|.$$

(9)

Here, n and p represent the numbers of data and parameters, respectively. Figure 3a shows examples of temporal changes in the difference in the cAIC of the two models for stations N.NSOV at Nasu, V.ASME at Asama, and V.KSHA at Kusatsu-Shirane. We defined the difference in cAIC (hereafter called "ΔcAIC") as
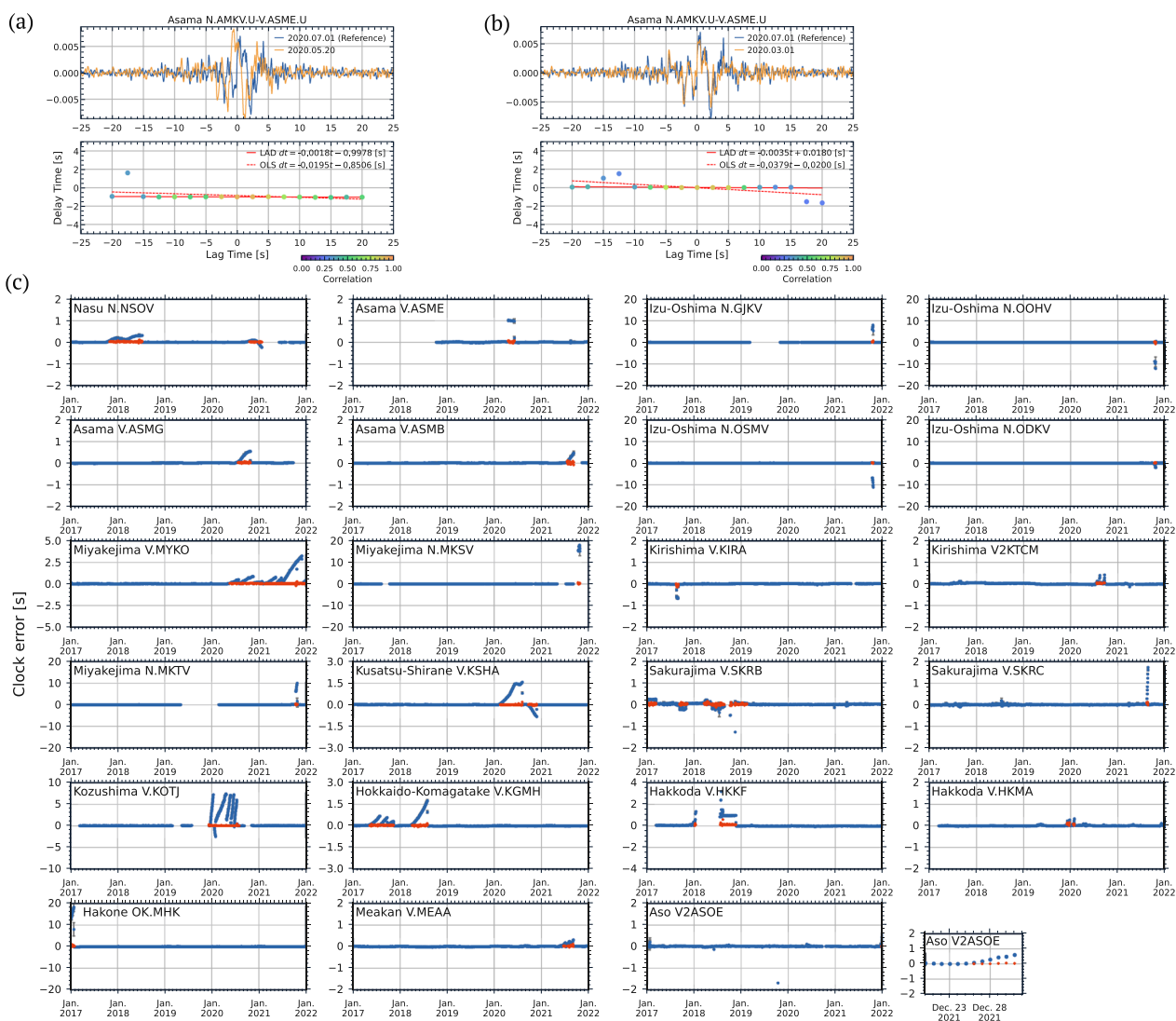
Fig. 2 Examples of the relative clock error estimation and estimated temporal changes in clock errors. **a** (top) Waveforms of the stacked cross-correlation functions (SCCFs) on days without clock errors (blue line) and on the day when the clock error occurred (orange line) for the station pair of N.AMKV–V.SAME at the Asama volcano. (bottom) Estimated delay times between the two SCCFs for each lag time. The solid and dashed red lines represent fitted straight lines using LAD and OLS regression, respectively. **b** Similar to **a**, but for the SCCFs when no clock error occurred. **c** Temporal changes in clock errors for 23 seismic stations at 13 volcanoes (blue circles). A positive clock error means that the clock of the station is delayed. The red crosses represent the clock errors estimated using SCCFs that were calculated after correcting timestamps of seismic ambient noise records

cAIC$(y = ax + b)$ − cAIC$(y = ax)$, hence, significantly small value of the ΔcAIC means that straight lines with intercept is better model. For all these three stations, the ΔcAIC become much smaller when clock errors occur. Additional file 1: Fig. S10a−c shows temporal changes in ΔcAIC at stations N.NSYV (Nasu), V.ASMO (Asama), V.ONNH (Ontake), respectively. Although estimated clock errors are sometimes highly scattered, ΔcAIC during these periods does not decrease significantly. This suggests that AIC will help us judge whether a clock error

occurs and judge when the clock error starts. It is worth noting that we found some cases in that ΔcAIC does not work well (Additional file 1: Fig. S10d, e). For these cases, large ΔcAIC fluctuations were observed even during periods when clock errors did not occur.

In this section, we have introduced two indicators that may be useful in judging whether the estimated clock error is significant. It will be important to combine several approaches: comparison of the estimated clock errors with the magnitude of background fluctuation, use

Table 1 Summary of estimation results of clock errors (see also Table 2). Note that stations in which clock errors occurred were not used for calculation of 98th percentiles of absolute values of clock errors

Volcano	Clock error (station, max. amplitude, period)	98th percentiles of clock error
Adatara	–	0.01 s
Akitakomagatake	–	0.01 s
Akitayakeyama	–	0.02 s
Aogashima	–	0.01 s
Asama	V.ASMB.U + 0.46 s (Jul.2021–Sep.2021)	0.03 s
	V.ASME.U + 1.02 s (Apr.2020–Jun.2020)	
	V.ASMG.U + 0.54 s (Jul.2020–Oct.2020)	
Aso	V2ASOE.U + 0.59 s (Dec.2021)	0.04 s
Atosanupuri	–	0.02 s
Azuma	–	0.01 s
Bandai	–	0.02 s
Chokaisan	–	0.07 s
Esan	–	0.01 s
Fuji	–	0.03 s
Hachijojima	–	0.01 s
Hakkoda	V.HKKF.U + 3.16 s (Jan.2018–Nov.2018)	0.01 s
	V.HKMA.U + 0.32 s (Dec.2019–Jan.2020)	
Hakone	OK.MHK.U + 18.27 s (Jan.2017–Mar.2017)	0.03 s
Hakusan	–	0.06 s
Hokkaido-Komagatake	V.KGMH.U + 0.67 s, + 1.75 s (May.2017–Nov.2017, Mar.2018–Aug.2018)	0.02 s
Ioto	–	0.02 s
Iwakisan	–	0.03 s
Iwatesan	–	0.01 s
Izu-Oshima	N.GJKV.U + 8.01 s (Oct.2021)	0.01 s
	N.ODKV.U -2.01 s (Oct.2021)	
	N.OOHV.U -12.22 s (Oct.2021)	
	N.OSMV.U -11.20 s (Oct.2021)	
Izu-Tobu	–	0.02 s
Kirishima	V.KIRA.U -0.68 s (Aug.2017)	0.02 s
	V2KTCM.U + 0.41 s (Jul.2020–Sep.2020)	
Kozushima	V.KOTJ.U + 7.42 s (Dec.2019–Jul.2020)	0.04 s
Kuchinoerabu	–	0.03 s
Kujyu	–	0.02 s
Kurikomayama	–	0.02 s
Kusatsu	V.KSHA.U + 1.56 s, -0.84 s (Feb.2020–Aug.2020, Sep.2020–Nov.2020)	0.01 s
Kuttara	–	0.02 s
Meakan	V.MEAA.U + 0.31 s (Jun.2021–Sep.2021)	0.02 s

of AIC, and re-estimation of the clock error after correcting the timestamp of seismograms.

Comparison with other estimation methods

To determine a suitable estimation method for clock errors in the 50 volcanoes, we performed the same analysis using the other two methods. The first method is the "WCC–OLS method." This is similar to the method used in this study (WCC–LAD method). Delay times at each lag time were first measured using the same short time window as the WCC–LAD method, and then a straight line was fitted using OLS regression. The second method is the "CC method" in which, a cross-correlation function between two SCCFs is computed without dividing

**Table 2** Summary of estimation results of clock errors

Volcano	Clock error (station, max. amplitude, period)	98th percentiles of clock error
Midagahara	–	0.02 s
Miyake	N.MKSV.U + 18.01 s (Oct.2021)	0.03 s
	N.MKTV.U + 9.98 s (Oct.2021)	
	V.MYKO.U + 3.23 s (May.2020–Dec.2021)	
Nasu	N.NSOV.U + 0.33 s, -0.24 s (Oct.2017–Jul.2018, Oct.2020–Jan.2021)	0.01 s
Niigatayakeyama	–	0.02 s
Niijima	–	0.01 s
Nikkoshirane	–	0.02 s
Norikuradake	–	0.01 s
Ontake	–	0.03 s
Sakurajima	V.SKRB.U -1.28 s (Jan.2017–Mar.2017, Sep.2017–Nov.2017, Mar.2018–Feb.2019)	0.02 s
	V.SKRC.U + 1.72 s (Aug.2021)	
Satsumaiwo	–	0.01 s
Suwanosejima	–	0.06 s
Taisetsuzan	–	0.02 s
Tarumae	–	0.02 s
Tokachi	–	0.02 s
Towada	–	0.02 s
Tsurumidake	–	0.01 s
Unzen	–	0.04 s
Usu	–	0.02 s
Yakedake	–	0.03 s
Zao	–	0.01 s

them into short time windows. The lag time at which the cross-correlation function between the two SCCFs reached a maximum corresponded to the relative clock error between the two stations. The maximum value of the cross-correlation function was searched between -5 and $+5$ s in lag time (between -10 and $+10$ s in lag time in the case of Kozushima, between -30 and $+30$ s in lag time in the cases of Miyakejima, Izu-Oshima, and Hakone) of this function.

Figure 3b shows examples of the temporal changes in clock errors estimated using three different methods: the WCC–LAD method (blue circles), the WCC–OLS method (green stars), and the CC method (red crosses) at Nasu, Asama, and Kusatsu-Shirane. The WCC–OLS method often underestimates clock errors. For example, clock errors for station V.ASME at Asama volcano from this method are about 0.1 s smaller than those from the other two methods between April and June 2020. Such an underestimation also occurred at Kusatsu-Shirane volcano. Because the OLS regression is more affected by outliers than the LAD regression (see also Fig. 2a, b), the WCC–OLS method is not suitable for reliable monitoring of clock errors at the 50 volcanoes. Compared with

the WCC–OLS method, the CC method yielded more stable results. The clock errors estimated using this method are almost identical to those from the WCC–LAD method in most cases (see small panels on the right side of Fig. 3b). However, clock errors estimated using the CC method are occasionally scattered. For example, clock errors at N.NSOV from the CC method are sometimes estimated to be around zero during October 2017−June 2018, although clock errors occurred during this period. Results of the WCC–LAD method did not show such a highly scattered clock error during the same period.

To discuss the possible causes of such scattered clock errors at station N.NSOV at the Nasu volcano, we compared the estimation results of the WCC–LAD and CC methods in more detail. The top panel of Fig. 3c shows the waveforms of the SCCF on January 3, 2020 and that on July 27, 2020, for station pairs N.NSOV–N.NSSV. SCCFs for these days have large amplitudes between -5 and $+5$ s in lag time, and their phases are clearly different along the lag time in this portion. The bottom left panel shows the delay times measured using the WCC–LAD method. Delay times of 2–5 s were estimated between -5 and $+5$ s in the lag time. However,
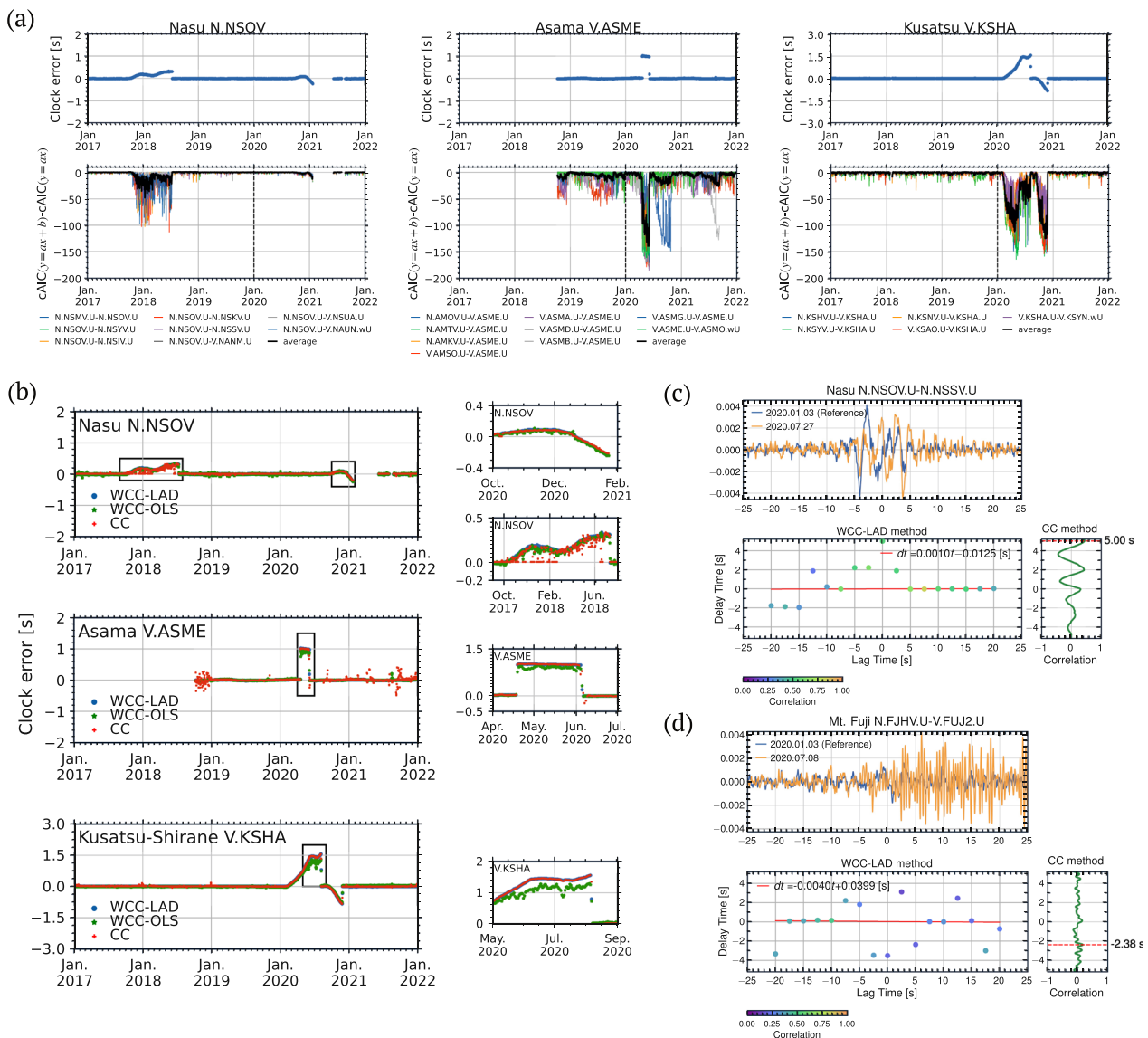
Fig. 3 Examples of temporal changes in the difference in the cAIC and comparison of clock error estimation results from different methods. **a** Temporal changes in clock errors (top) and those for differences in cAIC (ΔcAIC) from two different models for line fitting ($y = ax + b$ and $y = ax$) (bottom). Values of ΔcAIC were calculated between SCCF on the day indicated by the black vertical dashed line and an SCCF on another day. **b** Temporal changes in clock errors estimated using three different methods. Here, *WCC–LAD* windowed cross-correlation and least absolute deviation regression, *WCC–OLS* windowed cross-correlation and ordinary linear least-squares regression, *CC* cross-correlation without dividing into short time windows. The enlarged results for the period indicated by the black rectangle are shown in the small panels on the right side of the figure. **c** (Top) Waveforms of the stacked cross-correlation functions (SCCF) in January 3, 2017 (blue line) and SCCF in July 27, 2020 (orange line), for the station pair N.NSOV–N.NSSV at the Nasu volcano. (bottom left) Estimated delay times between the two SCCFs for each lag time and the fitted straight line with the least absolute deviation (LAD) regression (red line). (bottom right) The cross-correlation function was calculated using the −25 to +25 s lag time portion of the two SCCFs. The horizontal dashed red line represents the delay time that maximizes the cross-correlation function. **d** Similar to the panel **b**, but for the station pair N.FJHV–V.FUJ2 at Mt. Fuji

the result of the LAD regression (red straight line) was not affected by these outliers, and the clock error was estimated to be −0.01. In contrast, the CC method estimated the clock error to be +5 s (bottom-right panel). The significant difference in phases between

the two SCCFs for the lag time of −5 to +5 s should have contaminated the clock error estimation using the CC method. Seasonal variations in noise sources might affect the waveforms of seismic ambient noise CCFs. In fact, the energy flux from stations N.NSSV to

N.NSOV is more predominant in the SCCF on January 3, 2020; the acausal (negative lag time) part has a larger amplitude than the causal (positive lag time) part in that SCCF. On the other hand, the energy flux from station N.NSOV to N.NSSV was more predominant in the SCCF on July 27, 2020; it had a large amplitude in the causal part.

We report another possible case in which the estimation of clock errors using the CC method can be biased. Figure 3d is similar to Fig. 3c, but the results for the station pair N.FJHV–V.FUJ2 at Mt. Fuji. The waveform in the causal portion of the SCCF on July 8, 2020, corresponding to the wave propagation from N.FJHV to V.FUJ2, changed significantly compared with the SCCF on January 3, 2020. These large changes in waveforms in the SCCF portion corresponding to wave propagation from N.FJHV to the other station began in June 2020 and continued through August. Station N.FJHV is the westernmost station in the observation network of Mt. Fuji, and these waveform changes have continued for several months, suggesting that large-scale construction work might have been underway during this period there. Despite these large waveform changes in the SCCF, the WCC–LAD method can estimate clock error using delay time values in the time window in which the waveform correlation between the reference and current SCCFs is relatively high (bottom left panel of Fig. 3d). Because a large waveform change occurs in the SCCF, the maximum value of the cross-correlation function calculated using the CC method is approximately 0.2, and the lag time at this time is estimated to be -2.38 s, which is significantly different from the value estimated using the WCC–LAD method (bottom right panel of Fig. 3d).

These examples show that the WCC–LAD method can be used to stably estimate the clock error. By estimating the delay times between two SCCFs for several short time windows (WCC), the effect on clock error estimation can be suppressed even when a part of the SCCF waveform changes significantly, such as owing to fluctuations in noise sources. Moreover, the LAD regression, which is robust to outliers, further improves the robustness of clock error estimations.

Conclusions

This study comprehensively estimated instrumental clock errors at 50 active volcanoes in Japan based on seismic interferometry. We presented a new method to estimate clock errors more stably using WCC–LAD. This method was applied to seismic ambient noise records at 50 volcanoes between January 2017 and December 2021, and we found clock errors at 23 stations in 13 volcanoes during the study period. The maximum amplitude of the clock errors ranged between 0.24 and 18.27 s. These values were significantly larger than the amplitude of the background fluctuation in the clock errors (<0.05 s). A comparison of different approaches for clock error estimations shows that our method using WCC and LAD regression can reduce the effect of waveform changes in seismic ambient noise CCFs, possibly because of noise source fluctuations in the estimation of clock errors. Because this approach uses seismic ambient noise data, the continuous monitoring of instrumental clocks is possible in various regions. Managing the time stamps of data using the method proposed in this study will contribute to data quality assurance and ensure the reliability of analysis using time information in various study regions.

Abbreviations

ADMM	Alternating direction method of multipliers
AIC	Akaike information criterion
CCF	Cross-correlation functions
GPS	Global positioning system
JMA	Japan Meteorological Agency
LAD	Least absolute deviation
NIED	National Research Institute for Earth Science and Disaster Resilience
OBS	Ocean-bottom seismometer
OLS	Ordinary linear least-squares
SCCFs	Stacked cross-correlation functions
SNR	Signal-to-noise ratio
RCCF	Reference SCCF
WCC	Windowed cross-correlation

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40623-023-01798-4.

Additional file 1: Table S1. Estimated clock errors for all stations. **Fig. S1.** (a) Topographic maps of Atosanupuri, Meakan, Taisetsuzan, Tokachi, Tarumae, Kuttara, Usu, Hokkaido-Komagatake, and Esan volcanoes. The red inverted triangles on the topographic maps represent seismic stations for which clock errors were detected. (b) Record sections of seismic ambient noise CCFs for those 9 volcanoes in the frequency band between 0.2 and 4 Hz. These CCFs were calculated by stacking daily CCFs during the study period. Note that the daily CCFs when clock errors occurred were discarded. **Fig. S2.** Similar to Fig. S1, but for Hakkoda, Towada, Iwakisan, Akitayakeyama, Mt. Iwate, Akitakomagatake, Chokaisan, Kurikomayama, and Zao volcanoes. **Fig. S3.** Similar to Fig. S1, but for Azuma, Adatara, Bandai, Nasu, Nikkoshirane, Kusatsu-Shirane, Asama, Mt. Fuji, and Hakone volcanoes. **Fig. S4.** Similar to Fig. S1, but for Izu-Tobu, Izu-Oshima, Niijima, Kozushima, Miyakejima, Hachijojima, Aogashima, and Ioto volcanoes. **Fig. S5.** Similar to Fig. S1, but for Niigatayakeyama, Midagahara, Yakedake, Norikuradake, Ontake, and Hakusan volcanoes. **Fig. S6.** Similar to Fig. S1, but for Tsurumidake, Kujyu, Aso, Unzen, Kirishima, Sakurajima, Satsumaiwo, Kuchinoerabu, and Suwanosejima volcanoes. **Fig. S7.** Temporal changes in SCCF (three-day-stacked CCF) waveforms for station pairs with clock errors (N.AMOV–V.ASMG, V.AMSO–V.ASMB, N.MKSV–V.MKJA, N.MKTV–V.MKJA, OK.KOM–OK.MHK, and V.MEAB–V.MEAA). **Fig. S8.** Similar to Fig. S7 but for N.GJKV–V.OSMA, N.ODKV–V.OSSN, N.OOHV–V.OSFT, N.OSMV–V.OSFT, V.KIRA–V.KIMH, and V2KTCM–V.KIKH. **Fig. S9.** Similar to Fig. S7 but for V.SKA2–V.SKRB, V.SKRC–V.SKRD, V.HKKF–V.HMA2, V.HKMA–V.HKTK, and N.ASIV–V2ASOE. **Fig. S10.** Temporal changes in clock errors (top) and

those for differences in cAIC from two different models for line fitting (y=ax+b and y=ax) (bottom) for six stations.

Author contributions
TH (corresponding author) performed the analyses and drafted the manuscript. HU supervised the conduct of this study. All authors have read and approved the manuscript.

Availability of data and materials
Continuous seismograms were obtained from the Data Management Center of NIED (https://hinetwww11.bosai.go.jp/auth/?LANG=en). The DEM data, ALOS Global Digital Surface Model "ALOS World 3D—30 m (AW3D30)," used to create a topographic map, are available at https://www.eorc.jaxa.jp/ALOS/en/dataset/aw3d30/aw3d30_e.htm.

Declarations

Competing interests
The authors declare no competing interest.

Author details
[1]National Research Institute for Earth Science and Disaster Resilience, Tennodai 3-1, Tsukuba, Ibaraki, Japan.

References
Bensen GD, Ritzwoller MH, Barmin MP, Levshin AL, Lin F, Moschetti MP, Shapiro NM, Yang Y (2007) Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements. Geophys J Int 169:1239–1260. https://doi.org/10.1111/j.1365-246X.2007.03374.x

Boyd S, Parikh N, Chu E, Peleato B, Eckstein J (2011) Distributed optimization and statistical learning via the alternating direction method of multipliers. Found Trends Mach Learn 3:1–122. https://doi.org/10.1561/2200000016

Brenguier F, Campillo M, Takeda T, Aoki Y, Shapiro NM, Briand X, Emoto K, Miyake H (2014) Mapping pressurized volcanic fluids from induced crustal seismic velocity drops. Science 345:80–82. https://doi.org/10.1126/science.1254073

Campillo M, Paul A (2003) Long-range correlations in the diffuse seismic coda. Science 299:547–549. https://doi.org/10.1126/science.1078551

Curtis A, Gerstoft P, Sato H, Snieder R, Wapenaar K (2006) Seismic interferometry—turning noise into signal. Lead Edge 25:1082–1092. https://doi.org/10.1190/1.2349814

Gouédard P, Seher T, McGuire JJ, Collins JA, van der Hilst RD (2014) Correction of ocean-bottom seismometer instrumental clock errors using ambient seismic noisecorrection of ocean-bottom seismometer instrumental clock errors using ambient seismic noise. Bull Seismol Soc Am 104:1276–1288. https://doi.org/10.1785/0120130157

Hable S, Sigloch K, Barruol G, Stahler SC, Hadziioannou C (2018) Clock errors in land and ocean bottom seismograms: high-accuracy estimates from multiple-component noise cross-correlations. Geophys J Int. https://doi.org/10.1093/gji/ggy236

Hirose T, Nakahara H, Nishimura T, Campillo M (2020) Locating spatial changes of seismic scattering property by sparse modeling of seismic ambient noise cross-correlation functions: application to the 2008 Iwate-Miyagi nairiku (M w 6.9), japan, earthquake. J Geophys Res. https://doi.org/10.1029/2019jb019307

Hurvich CM, Tsai C-L (1990) Model selection for least absolute deviations regression in small samples. Stat Probab Lett 9:259–265. https://doi.org/10.1016/0167-7152(90)90065-F

Le BM, Yang T, Chen YJ, Yao H (2017) Correction of OBS clock errors using Scholte waves retrieved from cross-correlating hydrophone recordings. Geophys J Int 212:891–899. https://doi.org/10.1093/gji/ggx449

Maeda T, Obara K, Furumura T, Saito T (2011) Interference of long-period seismic wavefield observed by the dense Hi-net array in Japan. J Geophys Res. https://doi.org/10.1029/2011jb008464

Moreau L, Stehly L, Boué P, Lu Y, Larose E, Campillo M (2017) Improving ambient noise correlation functions with an SVD-based Wiener filter. Geophys J Int 211:418–426. https://doi.org/10.1093/gji/ggx306

Poupinet G, Ellsworth WL, Frechet J (1984) Monitoring velocity variations in the crust using earthquake doublets: an application to the Calaveras Fault, California. J Geophys Res 89:5719–5731. https://doi.org/10.1029/jb089ib07p05719

Seats KJ, Lawrence JF, Prieto GA (2012) Improved ambient noise correlation functions using Welch's method. Geophys J Int 188:513–523. https://doi.org/10.1111/j.1365-246X.2011.05263.x

Sens-Schonfelder C (2008) Synchronizing seismic networks with ambient noise. Geophys J Int 174:966–970. https://doi.org/10.1111/j.1365-246X.2008.03842.x

Stehly L, Campillo M, Shapiro NM (2007) Traveltime measurements from noise correlation: stability and detection of instrumental time-shifts. Geophys J Int 171:223–230. https://doi.org/10.1111/j.1365-246X.2007.03492.x

Takeo A, Forsyth DW, Weeraratne DS, Nishida K (2014) Estimation of azimuthal anisotropy in the NW Pacific from seismic ambient noise in seafloor records. Geophys J Int 199:11–22. https://doi.org/10.1093/gji/ggu240

Viens L, Van Houtte C (2020) Denoising ambient seismic field correlation functions with convolutional autoencoders. Geophys J Int 220:1521–1535. https://doi.org/10.1093/gji/ggz509

Wang Q-Y, Brenguier F, Campillo M, Lecointre A, Takeda T, Aoki Y (2017) Seasonal crustal seismic velocity changes throughout Japan. J Geophys Res 122:7987–8002. https://doi.org/10.1002/2017jb014307

Wu CFJ (1986) Jackknife, bootstrap and other resampling methods in regression analysis. Ann Statist 14:1261–1295. https://doi.org/10.1214/aos/1176350142