JGR Solid Earth

RESEARCH ARTICLE
10.1029/2023JB027413

Physical Mechanism for a Temporal Decrease of the Gutenberg-Richter *b*-Value Prior to a Large Earthquake

Ryo Ito[1] and **Yoshihiro Kaneko**[1]

[1]Graduate School of Science, Kyoto University, Kyoto, Japan

Key Points:

- Temporal decrease of a *b*-value is found in continuum models of rate-and-state faults with frictional heterogeneities
- Change in a *b*-value is caused by the reduction of rupture barrier effectiveness due to increasing shear stresses within creeping regions
- Aseismic slip facilitated by frictional heterogeneities is a key process controlling the temporal decrease of a *b*-value

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

Y. Kaneko,
kaneko.yoshihiro.4e@kyoto-u.ac.jp

Citation:

Ito, R., & Kaneko, Y. (2023). Physical mechanism for a temporal decrease of the Gutenberg-Richter *b*-value prior to a large earthquake. *Journal of Geophysical Research: Solid Earth*, *128*, e2023JB027413. https://doi.org/10.1029/2023JB027413

Received 6 JUL 2023
Accepted 2 DEC 2023

Author Contributions:

Conceptualization: Yoshihiro Kaneko
Data curation: Ryo Ito, Yoshihiro Kaneko
Formal analysis: Ryo Ito, Yoshihiro Kaneko
Funding acquisition: Yoshihiro Kaneko
Investigation: Ryo Ito, Yoshihiro Kaneko
Methodology: Ryo Ito, Yoshihiro Kaneko
Project Administration: Yoshihiro Kaneko
Resources: Yoshihiro Kaneko
Software: Ryo Ito, Yoshihiro Kaneko
Supervision: Yoshihiro Kaneko
Validation: Ryo Ito, Yoshihiro Kaneko
Visualization: Ryo Ito

Abstract Observations of seismicity prior to large earthquakes show that the slope of a Gutenberg-Richter magnitude-frequency relation, referred to as a *b*-value, sometimes decreases with time to the mainshock. Yet, underlying physical processes associated with the temporal change of a *b*-value remain unclear. Here we utilize continuum models of fully dynamic earthquake cycles with fault frictional heterogeneities and aim to simulate the temporal variation of a *b*-value. We first identify a parameter regime in which the model gives rise to an active and accelerating foreshock behavior prior to the mainshock. We then focus on the spatio-temporal pattern of the simulated foreshocks and analyze their statistics. We find that the *b*-value of simulated foreshocks decreases with time prior to the mainshock. A marked decrease in the resulting *b*-value occurs over the duration of less than a few percent of the mainshock recurrence interval, broadly consistent with foreshock behaviors and *b*-value changes as observed in nature and laboratory, rock-friction experiments. In this model, increased shear stresses on creeping (or velocity-strengthening) fault patches resulting from numerous foreshocks make these creeping patches more susceptible to future coseismic slip, increasing the likelihood of large ruptures and leading to a smaller *b*-value with time. This mechanism differs from a widely invoked idea that the decrease of a *b*-value is caused by a rapid increase in shear stress that promotes micro-crack growth, and offers a new interpretation of *b*-value changes prior to a large earthquake.

Plain Language Summary Foreshocks are earthquakes that occur before a large earthquake (i.e., a mainshock) and are related to the mainshock in both time and space, but the mechanism of foreshock generation is not well understood. Analysis of foreshock statistics for many large earthquakes showed that larger foreshocks tend to occur more frequently in the time leading to the mainshock, a phenomenon known as the temporal decrease of a Gutenberg-Richter *b*-value. For example, the temporal decrease of a *b*-value in the hypocentral region of the 2011 M_w 9.0 Tohoku-Oki earthquake was observed prior to the mainshock, indicating the potential usefulness of *b*-values for earthquake forecasting. However, the physical mechanism behind such observations is still unclear. Here, we develop a physics-based model of fault slip evolution and successfully simulate a temporal decrease in the *b*-value of foreshocks. In this model, an increased stress level in slowly slipping parts of a fault facilitates earthquake rupture growth, resulting in the temporal decrease of a *b*-value prior to the mainshock. The model results, which are broadly consistent with foreshock behaviors and *b*-value changes as observed in nature and laboratory experiments, have important implications for the mechanism of foreshock generation and earthquake forecasting.

1. Introduction

While large earthquakes preceded by foreshock activity are relatively uncommon, observational studies reported complex evolution of foreshock sequences prior to some moderate to large earthquakes (e.g., Ando & Imanishi, 2011; Bouchon et al., 2011; Brodsky & Lay, 2014; Dodge et al., 1996; Ellsworth & Bulut, 2018; Gulia et al., 2016; Kato et al., 2012; Schurr et al., 2014; Simon et al., 2021; Yoon et al., 2019). One of the well-known examples is from the 2011 M_w 9.0 Tohoku-Oki earthquake, where active and migrating foreshocks had started several weeks before the mainshock (Ando & Imanishi, 2011; Kato et al., 2012). Some of these foreshocks may be related to the occurrence of slow slip events (Y. Ito et al., 2013). However, it remains unclear what causes the spatio-temporal evolution of these foreshocks as well as those observed prior to other moderate to large earthquakes.

Understanding the mechanism of foreshock generation is also an important issue for earthquake forecasting. One of the standard ways to quantify foreshock activities is the estimation of a b_{GR}-value that corresponds to a slope of Gutenberg-Richter (GR) magnitude-frequency relation (Gutenberg & Richter, 1949):

Writing – original draft: Ryo Ito
Writing – review & editing: Ryo Ito, Yoshihiro Kaneko

$$\log_{10} N = a_{GR} - b_{GR} M \quad (1)$$

where N is the number of earthquakes greater than or equal to magnitude M, and a_{GR} is a measure of seismic activity. Parameter b_{GR} measures the relative number of large events compared to small events, and at $b_{GR} = 1.0$, the number of events increases by a factor of 10 for every unit decrease of event magnitude. Smaller b_{GR} means that the relative proportion of large events increases, and vice versa. Based on the monitoring of b_{GR} changes in time, several studies attempted to forecast whether a mainshock would follow soon after a large foreshock (e.g., Gulia & Wiemer, 2019). However, b_{GR}-based earthquake forecasting that can accurately estimate the location, magnitude, and timing of a future large earthquake remains difficult (e.g., Dascher-Cousineau et al., 2020), in part due to the lack of understanding of the physical mechanism associated with b_{GR} changes.

Previous analysis of foreshocks found that b_{GR} decreases with time before the occurrence of some moderate to larger earthquakes (Table 1) (Gulia et al., 2016; Nanjo et al., 2012; Papadopoulos et al., 2010; Schurr et al., 2014; Shimbaru & Yoshida, 2021; Simon et al., 2021). The temporal decrease of b_{GR} means that larger events occur more frequently toward the time of a mainshock. Nanjo et al. (2012) reported a decadal scale temporal decrease of b_{GR} in the eventual hypocentral region of the 2011 M_w 9.0 Tohoku-Oki earthquake. Papadopoulos et al. (2010) and Gulia et al. (2016) found that b_{GR} had started to decrease weeks before the 2009 M_w 6.3 L'Aquila earthquake (Italy). Gulia et al. (2018) reported that normalized and stacked b_{GR} obtained from seismicity prior to 31 large earthquakes had decreased 10% during the months to days prior to the mainshocks. Schurr et al. (2014) showed that b_{GR} had decreased over 3 years prior to the 2014 M_w 8.1 Iquique earthquake (Chile). Shimbaru and Yoshida (2021) analyzed the foreshock statistics of 30 earthquakes of M_w 6 or greater in Japan and found that b_{GR} had decreased over weekly or annual scales near the epicenter in 9 of them. Hence, these studies suggest that the temporal scale over which b_{GR} changes ranges from weeks to years.

Foreshocks have also been studied in laboratory experiments (e.g., Bolton et al., 2021; Goebel et al., 2013; Marty et al., 2023; McLaskey, 2019; McLaskey & Lockner, 2014; Scholz, 1968; Yamashita et al., 2021). Scholz (1968) observed a temporal change of b_{GR} of microfractures in rock deformation experiments, and found that b_{GR} depends primarily on the differential stress. Goebel et al. (2013) showed that b_{GR} and differential stress have an inverse relationship and that the amount of increase in differential stress can be estimated from the amount of a b_{GR} decrease. Bolton et al. (2021) showed that the reduction of b_{GR} is due not only to the increase of stress with time but also to increases in the fault slip velocity and loading rate. More recently, Yamashita et al. (2021) reported that in laboratory experiments with a 2-m long, heterogeneous fault with pre-existing gauge, a temporal decrease of b_{GR} was observed starting from a few seconds before the onset of a large slip event. Hence the temporal changes of b_{GR} are observed from laboratory scales to those of large earthquakes despite many orders of magnitude differences in spatial and temporal scales.

Several mechanisms have been proposed to explain the changes of b_{GR}. Temporal decrease of b_{GR} might manifest an increase in shear (or differential) stress in the epicentral region of a mainshock (Scholz, 1968, 2015). The greater the shear stress, the faster the growth of micro-cracks, forming larger fracture cores and leading to larger foreshocks (Scholz, 1968, 2015). However, given the long (>100 years) recurrence interval of large earthquakes and slow tectonic loading, it is unclear as to why the shear stress rapidly increases in the epicentral region of a mainshock in a relatively short period of time (less than a few percent of the recurrence interval). Recently, Bolton et al. (2021) proposed that the evolution of fault zone porosity might be an important factor in foreshock generation as it affects the slip velocity. In addition, a greater reduction of b_{GR} for faster slip velocity or loading rate in laboratory experiments (Bolton et al., 2021) implies that faults prone to creep (e.g., plate-boundary faults) (Harris, 2017; Schwartz & Rokosky, 2007) may tend to exhibit the temporal change of b_{GR}. Nevertheless, the mechanism of the temporal change of b_{GR}, especially in relation to fault creep, is not well understood.

Several continuum models of earthquake cycles based on rate and state friction have been used to explain active and accelerating foreshock behaviors (Beall et al., 2022; Cattania & Segall, 2021; Dublanchet, 2022; Dublanchet et al., 2013; Luo & Ampuero, 2018; Schaal & Lapusta, 2019; Skarbek et al., 2012; Yabe & Ide, 2017, 2018). Some of these studies consider models with frictional heterogeneities that produce active and accelerating foreshock behaviors (Dublanchet, 2020; Dublanchet et al., 2013; Luo & Ampuero, 2018; Schaal & Lapusta, 2019; Skarbek et al., 2012; Yabe & Ide, 2017, 2018). In particular, Yabe and Ide (2018) showed that a quasi-dynamic model of earthquake cycles with frictional heterogeneities can lead to inverse Omori's laws (i.e., the rate of foreshocks increases in time to the mainshock as a power law: t^{-n}, where t is time to the mainshock and n is the

Table 1
Examples of the Temporal Change of b_{GR} Values Prior to a Mainshock

Earthquake	Year	M_w	Change in b_{GR} value	Time scale	Reference
L'Aquila	2009	6.3	1.09 → 0.68	~10 days	Papadopoulos et al. (2010)
Tohoku-Oki	2011	9.0	0.74 → 0.47	~10 years	Nanjo et al. (2012)
Iquique	2014	8.1	0.75 → 0.6	~3 years	Schurr et al. (2014)
Lab earthquakes	2021	−4	0.53 ± 0.09 → 0.15 ± 0.03	~3 s	Yamashita et al. (2021)

exponent). Dublanchet (2020) examined the dependence of normal stress and critical nucleation size on b_{GR}. The change of b_{GR} in relation to stress evolution over an earthquake cycle was examined by Dublanchet (2022) and Beall et al. (2022). However, it is not yet clear how these results are related to the temporal change of b_{GR} as observed in nature and laboratory rock-friction experiments, with the duration of less than a few percent of presumed mainshock recurrence intervals (Table 1).

In this study, we develop fully dynamic models of earthquake cycles with frictional heterogeneities and aim to simulate a temporal variation of b_{GR} before a mainshock. In Section 2, we describe our modeling approach and assumptions. In Section 3, we identify a parameter range in which active and accelerating foreshocks occur. In Section 4, we present a representative model in which the temporal decrease of b_{GR} occurs and explain the physical mechanism. We further explore the dependence of b_{GR} changes on model parameters (Section 5) and discuss the implications (Section 6).

2. Model Setup

We envision a situation in which a fault surface embedded into Earth's crust is characterized by a heterogeneous distribution of friction properties (Figures 1a and 1b). For numerical efficiency, we assume a one-dimensional, planar fault embedded into an infinite, elastic homogeneous continuum and simulate earthquake cycles subject to slow, tectonic loading. Our approach is based on a boundary integral method (Lapusta & Liu, 2009; Lapusta et al., 2000), adapted for a two-dimensional (in-plane, Mode II), fully dynamic model of earthquake cycles (Kaneko et al., 2016, 2017). The relation between fault slip velocity $V(x, t)$ and the corresponding shear stress $\tau(x, t)$ is given by

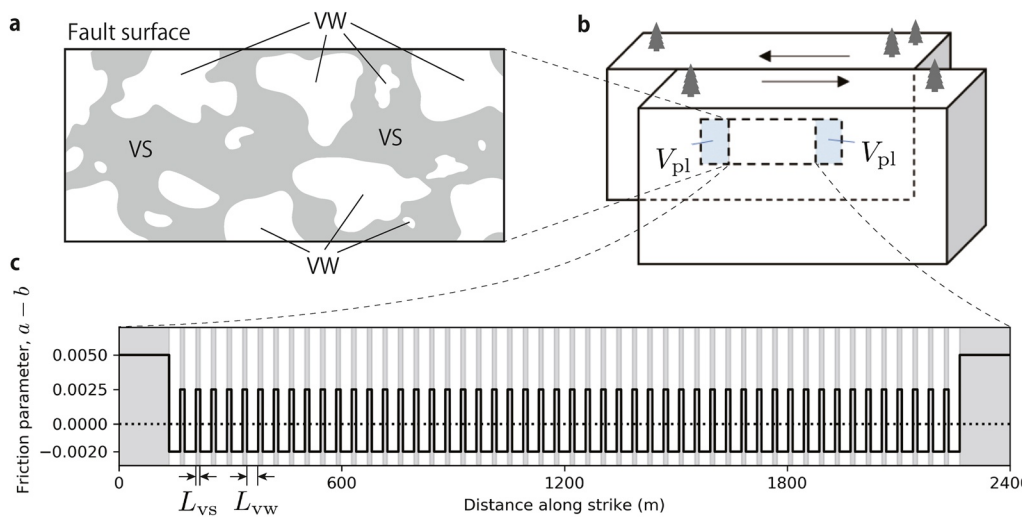

Figure 1. Fault model with frictional heterogeneities. (a) A conceptual illustration of a two-dimensional fault cross-section with frictional heterogeneities. Velocity-strengthening (VS, gray area) patches and velocity-weakening (VW, white area) patches are irregularly distributed on a fault surface. (b) Cartoon illustrating a strike-slip fault in the Earth's crust. For simplicity and numerical efficiency, we assume a one-dimensional fault. A uniform plate velocity $V_{pl} = 2$ mm/yr is applied to the outer 1,200 m at both ends of the fault. (c) Spatial distribution of rate-and-state frictional parameters $a - b$ on the fault, with alternating VS and VW patches. In the VS patches, $a - b$ is positive, and in the VW patches, $a - b$ is negative. The lengths of each VS patch $L_{vs} = 12$ m, and the length of a VW patch $L_{vw} = 30$ m.

Table 2
Model Parameters Assumed in This Study

	Parameter	
P	Density	2,610 kg/m^3
μ	Rigidity	23.3 GPa
N	Poisson's ratio	0.216
V_{pl}	Plate loading rate	2.0 mm/yr
$\dot{\tau}$	Shear stress loading uniformly added to the entire fault	0.020 MPa/yr
Σ	Effective normal stress	50 MPa
f_0	Reference friction coefficient	0.30–0.60
D_{RS}	Characteristic slip for state evolution	15 μm
c_p	P-wave speed	5.0 km/s
c_s	S-wave speed	3.0 km/s
a_{vs}	Rate-and-state friction parameter a in VS patches	0.020
a_{vw}	Rate-and-state friction parameter a in VW patches	0.020
b_{vs}	Rate-and-state friction parameter b in VS patches	0.016–0.019
b_{vw}	Rate-and-state friction parameter b in VW patches	0.021–0.024
L_{vs}	Length of a VS patch	12 m
L_{vw}	Length of a VW patch	30 m

$$\tau(x,t) = \tau^o(x) + f(x,t) - \frac{\mu}{2c_s}V(x,t) \tag{2}$$

where μ is the shear modulus, c_s is the shear wave speed, τ^o is the loading stress that would act on the interface if it were constrained against any slip, and $f(x,t)$ is a linear functional of prior slip over the causality cone (Lapusta & Liu, 2009; Lapusta et al., 2000). Note that, unlike widely used quasi-dynamic models (Dublanchet, 2022; Dublanchet et al., 2013; Luo & Ampuero, 2018; Skarbek et al., 2012; Yabe & Ide, 2017, 2018), our approach is fully dynamic and accounts for full inertial effects, which would be important for simulating rupture growth and quantifying the eventual sizes of earthquakes.

The constitutive response of the fault is governed by rate-and-state friction laws with the aging form of state variable evolution (Dieterich, 1978, 1979; Ruina, 1983):

$$\tau = \sigma\left[f_0 + a\ln\left(\frac{V}{V_0}\right) + b\ln\left(\frac{V_0\theta}{D_{RS}}\right)\right] \tag{3}$$

$$\frac{d\theta}{dt} = 1 - \frac{V\theta}{D_{RS}} \tag{4}$$

where τ is the shear strength, σ is the effective normal stress, a and b are the rate and state constitutive parameters, V is slip rate, f_0 is the reference friction coefficient at $V = V_0$, θ is a state variable, and D_{RS} is the characteristic slip for state variable evolution.

We assume relatively simple distributions of fault frictional heterogeneities. Alternating patches of velocity strengthening (VS, $a - b > 0$) and velocity weakening (VW, $a - b < 0$) are assumed (Figure 1c). The fault length is 2,400 m, and there are 1,200 m regions at both ends of the fault (not shown in Figure 1c) that are loaded by imposed loading rate $V_{pl} = 2$ mm/yr. To avoid loading only from the sides of the fault, we additionally apply a time-independent stressing rate $\dot{\tau} = 0.02$ MPa/yr uniformly along the fault. Without the uniform stressing rate, the fault is always gradually unzipped from the edge toward the center, as a result of the assumption that the fault is 1D. The calculated number of cells is 16,384 over the 2,400 m fault, with each computational cell size $\Delta x = 0.15$ m. The number of VS patches (excluding both ends of the fault) is 50 and the number of VW patches is 51, and the length of the VS and VW patches $L_{vs} = 12$ m and $L_{vw} = 30$ m, respectively. We set a uniform value of $D_{RS} = 15$ μm, which is within a range of typical values reported in rock-friction laboratory experiments (Ikari, Niemeijer, & Marone, 2011; Marone, 1998). Model parameters and their range explored in this study are summarized in Table 2.

Under slow tectonic loading, a stick-slip frictional instability (i.e., unstable slip) can develop only if the rate-weakening region of the fault exceeds the nucleation length h^* (e.g., Dieterich, 1992; Rice, 1993). Two theoretical estimates for a nucleation size relevant to our model are

$$h^*_{RA} = \frac{2}{\pi}\frac{\mu' b_{vw} D_{RS}}{\sigma(b_{vw} - a_{vw})^2} \tag{5}$$

(Rubin & Ampuero, 2005) and

$$h^*_{RR} = \frac{\pi}{4}\frac{\mu' D_{RS}}{\sigma(b_{vw} - a_{vw})} \tag{6}$$

(Rice, 1993; Ruina, 1983), where σ is the normal stress, $\mu' = \mu/(1-\nu)$, μ is shear modulus, and ν is Poisson's ratio. h^*_{RR} was derived from the linear stability analysis of steady sliding (Rice, 1993; Ruina, 1983), whereas h^*_{RA} is valid for $a/b \gtrsim 0.5$ and was derived based on energy balance for a quasi-statically expanding crack (Rubin & Ampuero, 2005). An actual nucleation size in simulations also depends on the background loading rate and how a VW patch is loaded by aseismic slip in the neighboring VS patches (Kaneko & Lapusta, 2008; Kaneko et al., 2016).

The accuracy of numerical solutions in fully dynamic earthquake cycle problems depends on the resolution of a cohesive zone, that is, the portion of the fault behind the crack tip where the shear stress decreases from its peak value to its dynamic value (Day et al., 2005). To properly resolve the cohesive zone, the ratio of cohesive zone size to the cell size Δx needs to be at least ∼3 (Kaneko et al., 2008). In our typical models, there are at least 3 cells within a cohesive zone to ensure that the numerical resolution is adequate. The total number of time steps in a typical simulation that produces ∼10 mainshock cycles is about 50–250 million, which takes about 3–20 days on a single Dell workstation (Precision Tower CTO Base, 4.8 GHz Intel Xeon W-1250P).

3. Parameter Exploration to Search for Accelerating Foreshock Behavior

We first vary the distributions of friction parameters $a_{vs} - b_{vs}$ and identify the cases where active foreshock sequences occur prior to mainshocks. Figure 2 shows the spatio-temporal evolution of seismicity in models with different degrees of $a_{vs} - b_{vs}$. We define the onset of an earthquake as when the maximum slip rate exceeds 1 cm/s and the end of an earthquake as when the maximum slip rate falls below 0.9 cm/s. Slightly different slip-rate thresholds are used to avoid double counting of earthquakes caused by slightly oscillating slip rates around the thresholds. We refer to slip that occurs below these slip-rate thresholds as "aseismic slip" or "creep." In all these cases, $h^*_{RA} = 31$ m and $h^*_{RR} = 3.5$ m in the VW patches. Since earthquakes occur in all these cases, the actual nucleation sizes in the simulations are smaller than the width of a VW patch (=30 m).

We calculate moment magnitude M_w of an simulated earthquake using the following relation (Kanamori, 1977):

$$M_w = \frac{2}{3}\left(\log_{10} M_0 - 9.1\right), \tag{7}$$

where $M_0 = \mu DA$ is a seismic moment measured in Newton·meter, μ is shear modulus, D is the amount of coseismic slip, and A is the size of the coseismic slipped area. Since the fault is 1D, we simply set $A = l^2$, where l is the coseismic rupture length. In Figure 2b, M_w of earthquakes range from ∼0 to ∼4.4.

When $a_{vs} - b_{vs}$ is small, the model leads to simple earthquake cycles with relatively few foreshocks (Figures 2a and 3a). Since earthquake rupture tends to propagate through weaker VS barriers, only a few small events occur 1 hour before a mainshock in this case (Figure 3a). In contrast, when $a_{vs} - b_{vs}$ is large, the VS barriers become more effective, and consequently, individual VW asperities tend to fail independently and irregularly (Figures 2c and 3c). As a result, swarm-like earthquake behavior with no discernible mainshock emerges (Figure 3c). Active and accelerating foreshock behavior occurs within a relatively narrow parameter regime in between the models with few foreshocks and earthquake swarms (Figures 2b and 3b). The amplitude of potency rates in the foreshocks increases with time (Figure 3b), indicating that larger events occur more frequently closer in time toward the mainshock. In the case of accelerating foreshocks (Figure 2b), the mainshock hypocenters tend to be near the center of the fault, but not in the case of earthquakes with few foreshocks (Figure 2a). The magnitude completeness is controlled by the critical nucleation size and the size of individual VW patches, where seismic events that rupture one VW patch correspond to $M_w \sim 2.1$. Dependence of seismicity patterns on the degree of frictional heterogeneities (i.e., $a_{vs} - b_{vs}$ in this case) is qualitatively consistent with previous studies that employed similar numerical models (Dublanchet, 2018; Dublanchet et al., 2013; Skarbek et al., 2012; Yabe & Ide, 2017, 2018). In the next section, we focus on the case in which accelerating foreshocks occur and analyze the foreshock statistics.

4. Representative Model for Accelerating Foreshocks Behavior

4.1. Spatio-Temporal Pattern of Simulated Foreshocks

In the case of accelerating foreshocks behavior, the spatio-temporal distribution of foreshocks is quite complex and qualitatively resembles those observed in nature and laboratory experiments. To visualize the evolution of accelerating foreshocks in detail, we zoom in on the final 24 hr prior to one of the mainshocks (Figure 4). When one foreshock ruptures a VW patch and arrests due to the surrounding VS patches, afterslip occurs within these VS patches, which increases shear stresses in the neighboring VW patches and can trigger the next foreshock. Successive foreshocks in the vicinity appear as a migrating foreshock pattern (Figure 4). The speeds of migrating foreshocks are of the order of km per day, comparable to those reported weeks before the M_w 9.0 Tohoku-Oki earthquake (Kato et al., 2012). The apparent migration speeds tend to increase in time toward the mainshock. The mainshock rupture eventually nucleates in the middle of the foreshock zone (Figure 4), and the foreshocks

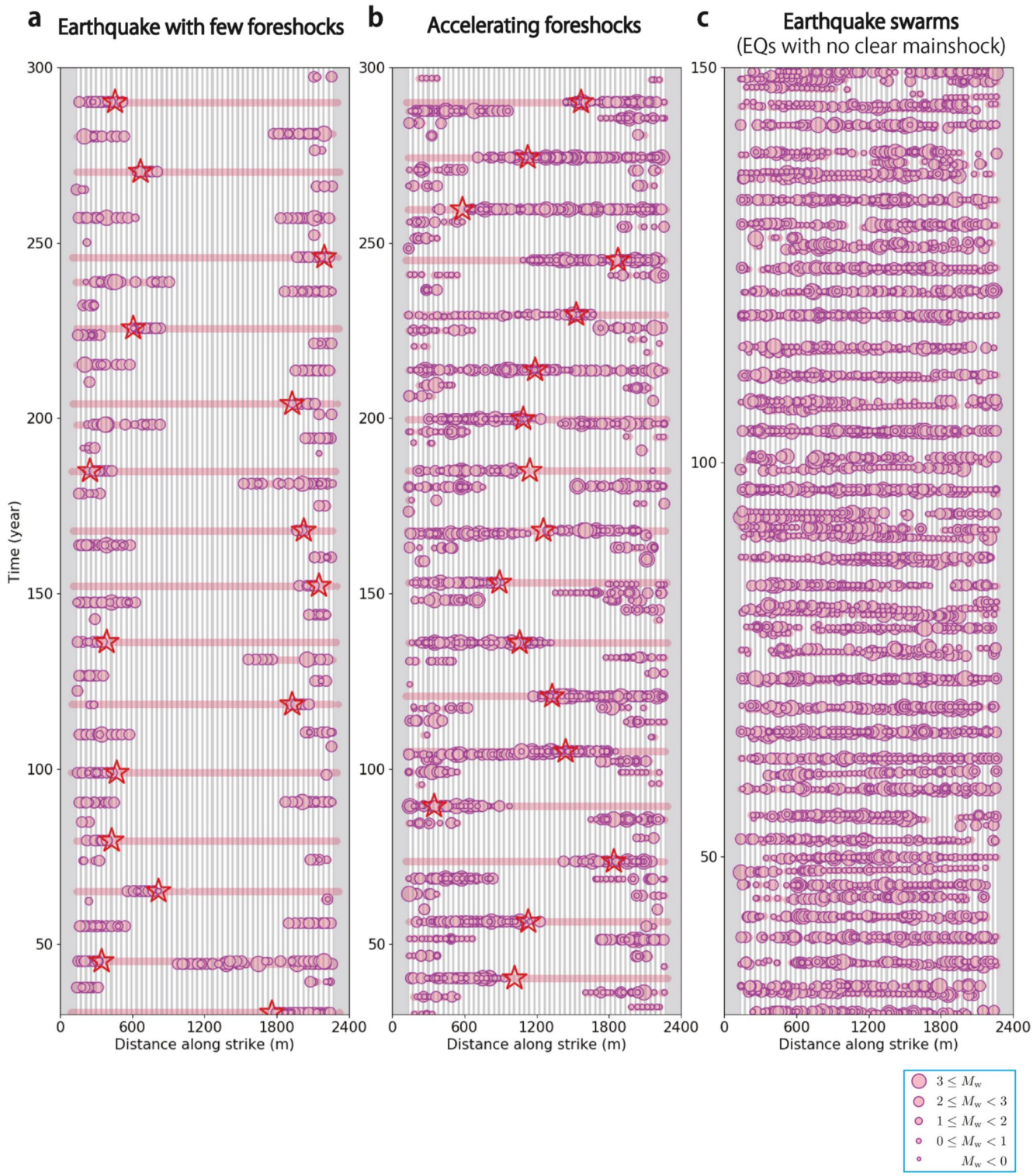

Figure 2. Seismicity patterns over multiple earthquake cycles in the cases with different degrees of frictional heterogeneities. Purple circles represent foreshock epicenters. Pink regions represent the rupture area of each event (plotted when slip rate >1 cm/s). Background gray and white areas represent velocity strengthening and velocity weakening (VW) patches, respectively. The red stars are the mainshocks that rupture the entire VW patches of the fault. (a) Case with mainshocks accompanying few foreshocks ($\sigma(a_{vs} - b_{vs}) = 0.050$ MPa). Quasi-periodic mainshocks happen every ∼19 years. The mainshock magnitude M_w is ∼4.5. (b) Case with mainshocks accompanying active foreshocks ($\sigma(a_{vs} - b_{vs}) = 0.125$ MPa). Quasi-periodic mainshocks happen every ∼15 years. The mainshock magnitude M_w is ∼4.4. (c) Seismicity pattern over multiple earthquake cycles in the earthquake swarm case ($\sigma(a_{vs} - b_{vs}) = 0.20$ MPa). The largest earthquake does not rupture the entire fault in this case.

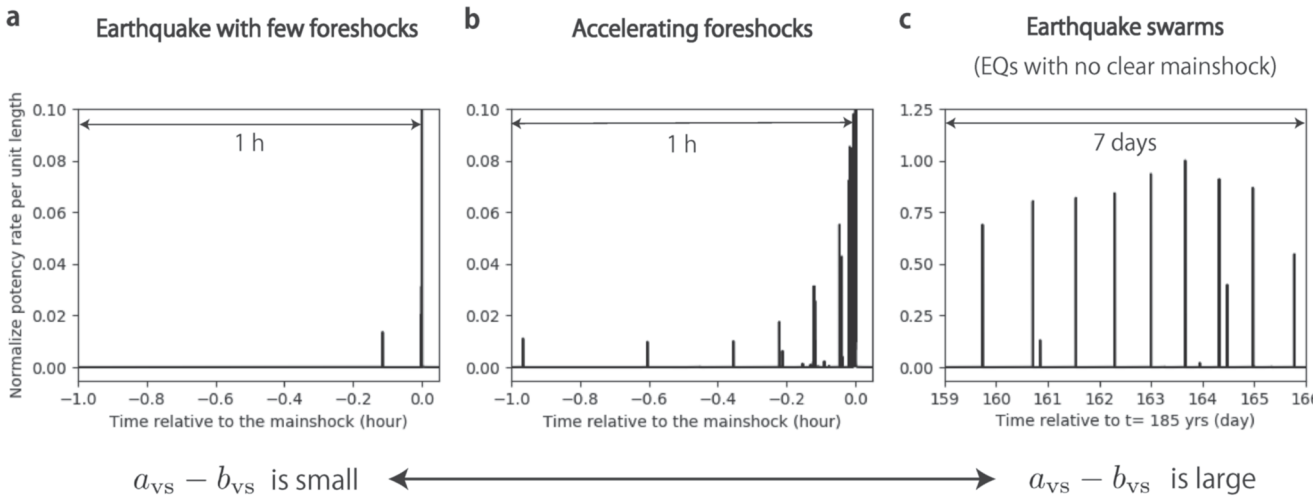

Figure 3. Different earthquake behaviors in the cases with different degrees of frictional heterogeneities. (a) Normalize potency rates before and after one of the mainshocks in the case shown in Figure 2a. The potency rate is normalized by the maximum value corresponding to the mainshock potency rate. Individual spikes represent earthquakes, with the largest one being the mainshock. A relatively small number of foreshocks ($\lesssim$10 within 1 hr before the mainshock) occur. (b) Normalize potency rates before and after one of the mainshocks in the case shown in Figure 2b. Active and accelerating foreshocks happen before the mainshock. (c) Normalize potency rates over 7 days in the case shown in Figure 2c. The model leads to earthquake swarm behavior with no clear mainshock.

become more active with time toward the mainshock, similar to those in the laboratory experiments of Yamashita et al. (2021). The nucleation sizes of a foreshock and a mainshock are similar, and the maximum size of a mainshock is limited by the overall extent of VW patches (i.e., the fault length).

Upon examining how each foreshock triggers the next foreshock, we find that the contribution of stress increase within a VW patch due to afterslip in the neighboring VS patches is generally larger than that of the static stress change from the previous foreshock (Figure 5 and Figure S1 in Supporting Information S1). Hence afterslip in the VS patches and increased loading from the afterslip tend to trigger the nucleation of a foreshock in this model, similar to the numerical result of Lui and Lapusta (2016). This means that aseismic slip occurring over relatively small regions (i.e., VS patches) is a key process controlling the foreshock evolution.

Dynamic triggering of earthquakes manifested as isolated seismic slip occurring at the coseismic time scale of the previous event can also be seen in this example, albeit happening less frequently. In several places, pink regions without purple circles (Figure 4) correspond to areas where seismic slip occurs at isolated places away from the hypocenter of foreshocks. Dynamically triggered events are relatively few and are not categorized as individual events due to the difficulty in defining them.

4.2. Temporal Evolution of b_{GR}

To examine the foreshock statistics, we plot the frequency-magnitude relations over different time periods in the case with accelerating foreshock behavior (Figure 6). To ensure that there are enough events for the analysis, the time windows are set cumulatively from the previous mainshock (Figure 6), and in each time window, all the foreshocks are summed over all mainshock cycles (Figure 2b). We find that the slopes estimated from the frequency-magnitude relations decrease with time, indicating the temporal decrease of b_{GR} (Figure 6).

We then calculate b_{GR} for different time windows using a maximum likelihood estimation method (Aki, 1965; Utsu, 1965) commonly used in observational studies:

$$b_{GR} = \frac{\log_{10} e}{\overline{M_w} - (M_c - \Delta M/2)} \tag{8}$$

where M_c is the magnitude of catalog completeness threshold (observational cut-off), $\overline{M_w}$ is the average magnitudes of all targeted foreshocks with a magnitude greater than M_c, and $\Delta M = 0.1$ is the magnitude interval. The uncertainty estimates are based on Shi and Bolt (1982).

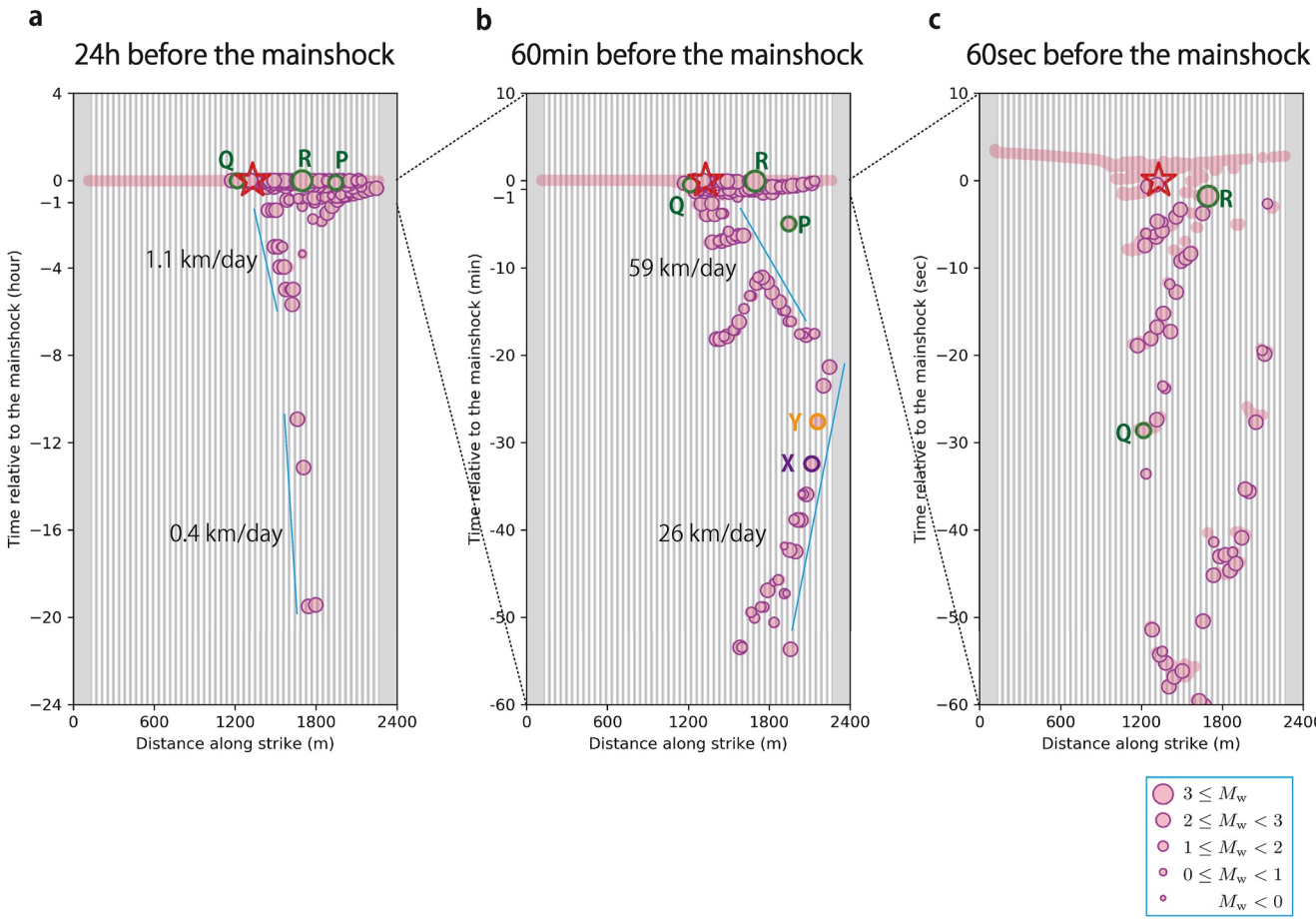

Figure 4. Spatio-temporal evolution of foreshocks in the last 24 hr leading to one of the mainshocks in Figure 2b. Foreshock distributions over the periods of (a) 24 hr, (b) 60 min, and (c) 60 s before the mainshock. Panel (b) is an enlarged image of panel (a) and panel (c) is an enlarged image of panel (b). Purple circles represent epicenters of foreshocks, and pink regions represent the rupture extent of each event (plotted when slip rate >1 cm/s). Background gray and white areas represent velocity strengthening and velocity weakening patches, respectively. The red star is the mainshock (M_w 4.4). Approximately 300 foreshocks occur in the final 24 hr before the mainshock. The green circles are Foreshocks "P," "Q," and "R" shown in Figure 9. The purple and orange circles are Foreshocks "X" and "Y" shown in Figure 5. Blue lines indicate the speeds of migrating foreshocks.

We find that estimated b_{GR} indeed decreases by ~30% over 1 month leading to the mainshock (the inset in Figure 6), meaning that larger foreshocks occur more frequently in the period closer to the mainshock. This behavior is qualitatively consistent with observations of b_{GR} decrease in nature (Table 1) (e.g., Gulia et al., 2016; Nanjo et al., 2012) and laboratory experiments (e.g., Bolton et al., 2021; Yamashita et al., 2021). However, the calculated b_{GR} ranges from 1.5 to 2.5, larger than typical values of ~1.0 often found in nature. This might be because the fault is 1D, making rupture arrest more likely when encountering a VS patch, whereas the rupture can go around a localized VS patch on a 2D fault. Hence, we focus on relative changes in b_{GR} in this study instead of the absolute values.

Figure 6 shows that the foreshock statistics do not strictly follow GR frequency-magnitude relations in that the slope is not a straight line. The rupture extent of $M_w \sim 2.1$ and $M_w \sim 2.7$ foreshocks roughly corresponds to one and two VW patches, respectively (Figure 6). Since the nucleation size in this model is only slightly smaller than a VW patch size, the major break in the slope at $M_w \sim 2.1$ is caused by the nucleation size in our model. For this reason, we set $M_c = 2.1$. Note that decreasing the nucleation size significantly increases the computational cost. In addition, a slight deviation from a straight line at $M_w \sim 2.7$ corresponds to events that rupture two VW patches (Figure 6). Therefore, the slight deviation from the straight line would be an artificial effect of the regular arrangement of the frictional heterogeneities assumed in this model. Despite these limitations, it is remarkable that the temporal decrease of b_{GR} is successfully simulated in this relatively simple model.

A more detailed temporal variation of b_{GR} further confirms that the b_{GR} continuously decreases with time (Figure 7). b_{GR} sharply decreases at ~1 hr before the mainshock when relatively large foreshocks occur around that time

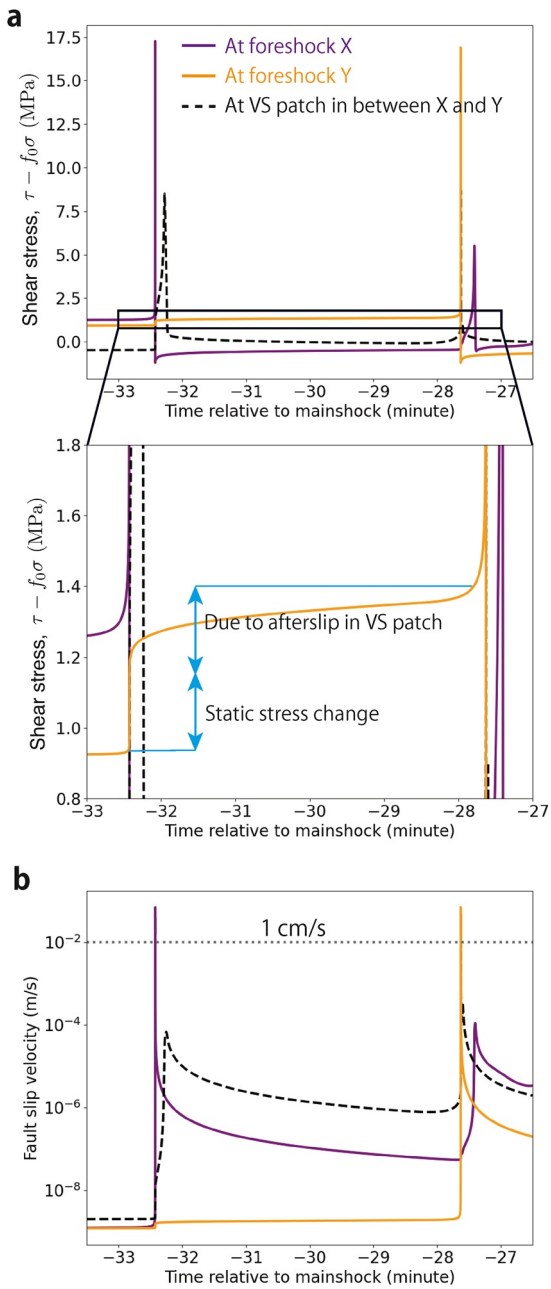

Figure 5. Stress changes induced by a foreshock to a neighboring velocity weakening (VW) patch. (a) Shear stress, $\tau - f_0\sigma$ and (b) slip velocity as a function of time at the center of the VW patches where Foreshocks "X" and "Y" (Figure 4b) occur, and at the center of the velocity strengthening (VS) patch (dashed black) in between the VW patches. We define the end of the static stress change as the time at the end of the previous foreshock plus the additional time for the shear waves to arrive at the relevant VW patch. In the triggering of the neighboring foreshock (Foreshock Y), the contribution of shears stress increase due to afterslip in the VS patch is larger than that of the static stress change from Foreshock X.

(Figure 7). We also check if the selection of time widows does not bias the result of b_{GR} changes. Instead of cumulative time windows, non-overlapping time windows are used to plot the frequency-magnitude relations (Figure S2 in Supporting Information S1). The result still shows the temporal decrease of b_{GR} in that case (Figure S2 in Supporting Information S1).

4.3. Why Does b_{GR} Decrease With Time Before a Mainshock?

To find the cause of the decrease in b_{GR}, we plot the distributions of the fault slip velocity and shear stress at certain times before the mainshock (Figure 8). Slip rates on the fault generally increase with time due to the afterslip of numerous foreshocks. Notably, the average shear stress on the entire fault remains nearly constant over this period (Figure 8c). As the duration of 1 month is much shorter than the recurrence period (∼17 years) of the mainshocks, the tectonic loading has a negligible effect during this time period. Foreshocks generally result in reducing the shear stresses within the ruptured VW patches. In contrast, the average shear stress over the VS patches increases with time (Figure 8c). Hence the temporal change in b_{GR} must be related to the increase of slip rates and shear stresses in the VS patches.

To understand how foreshock rupture grows large, we examine the source characteristics of individual foreshocks of different magnitudes. Figure 9 shows representative foreshocks of different magnitudes in the case shown in Figure 4. For a small foreshock (M_w 2.0), the rupture is nucleated in a VW patch and arrests in the neighboring VS patches (Figure 9a). The shear stresses in the ruptured VW patch decrease while they increase in the neighboring VS patches, as expected. During a larger foreshock, the rupture propagates through multiple VS patches where the shear stress level before the event is already fairly large due to the effect of previous foreshocks (e.g., VS1 and VS2 in Figure 9b). In other words, these VS patches that are slipping aseismically before the start of the foreshock promote earthquake rupture to propagate through the VS patches (Figures 9b and 9c). These VS patches that are ruptured co-seismically experience a slight shear-stress increase, which might be counterintuitive (e.g., VS1 and VS2 in Figure 9b). Foreshock rupture tends to arrest when it encounters the VS regions in which shear stress before the rupture is relatively low. Some of the VW patches slipped co-seismically during a large foreshock experience a small stress decrease (≲0.3 MPa) (e.g., VW2 and VW3 in Figure 9b). This happens because co-seismic slip in previous foreshocks has led to a decrease in shear stress in the VW patches that are not yet healed at the time of the next foreshock. Hence, seismic slip repeatedly occurs on the same VW patches (Figure 4). Note that the static stress drop of a foreshock, which is calculated from the shear stress changes averaged over the slip area defined as the area of slip greater than 20% of the maximum slip, does not depend on its magnitude. For example, the static stress drop of $M_w = 2.0$, 2.9, and 3.1 events (Figure 9) is 0.74, 1.2, and 0.75 MPa, respectively. From this analysis, a larger level of shear stresses in the VS patches leads to larger foreshocks statistically, which in turn, makes b_{GR} smaller.

To establish a more quantitative understanding of how the rupture of a foreshock grows, we follow the analysis of Kaneko et al. (2010) and consider an idealized situation on a rate-and-state fault under steady-state conditions ($d\theta/dt = 0$).

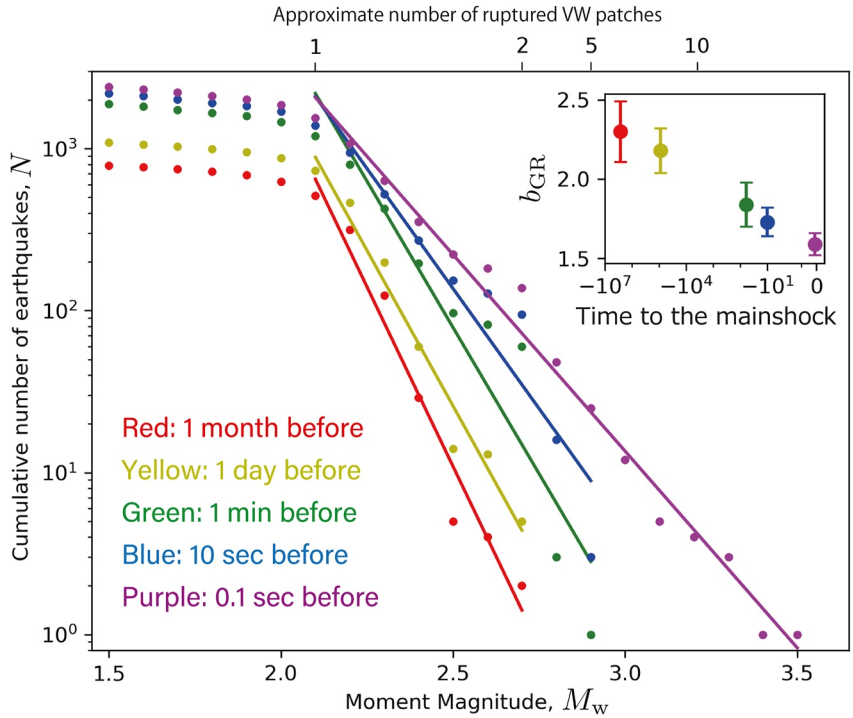

Figure 6. Temporal change of b_{GR} in the case shown in Figure 2b. Relationship between magnitude M_w and cumulative number of earthquakes N are plotted over five-time windows: from the previous mainshock until X before the next mainshock with X being 1 month (red), 1 day (yellow), 1 min (green), 10 s (blue), and 0.1 s (purple). All the foreshocks over multiple mainshock cycles shown in Figure 2b are included. The approximate number of velocity weakening patches ruptured in different magnitude events is also indicated in the top axis. The inset figure shows the evolution of b_{GR} with time in the log scale. b_{GR} is obtained by the maximum likelihood estimation (Equation 8) with $M_c = 2.1$. The fitting lines are obtained by the least squares method. b_{GR} decreases with time toward the mainshock.

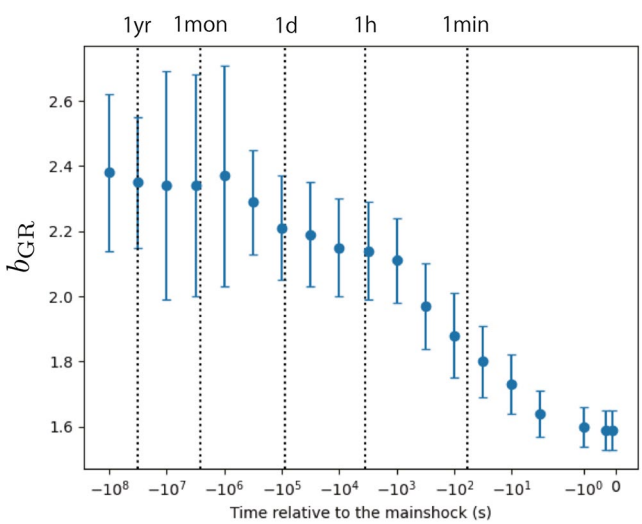

Figure 7. Temporal variation of b_{GR} in the case shown in Figure 2b. b_{GR} is calculated in the time window from the previous mainshock to the time shown in the horizontal axis. b_{GR} continuously decreases with time toward the mainshock, with the rate being significantly faster starting from 1 hr before the mainshock.

Under such conditions, the state variable θ can be expressed as $\theta = D_{RS}/V$ from Equation 3 and shear stress in VS patches τ_{vs} can be approximated as

$$\tau_{vs} = \sigma\left[f_0 + (a_{vs} - b_{vs})\ln\left(\frac{V_{vs}}{V_0}\right)\right] \quad (9)$$

where V_{vs} is slip velocity in the VS patches. During seismic slip with slip rate V_{vs}^{dyn}, shear stress in the VS patches can be approximated as

$$\tau_{vs}^{dyn} = \sigma\left[f_0 + (a_{vs} - b_{vs})\ln\left(\frac{V_{vs}^{dyn}}{V_0}\right)\right]. \quad (10)$$

Hence a shear stress increase required to reach the seismic slip rate in a VS patch $\Delta\tau_{vs}$ can be approximated as

$$\Delta\tau_{vs} = \tau_{vs}^{dyn} - \tau_{vs} = \sigma(a_{vs} - b_{vs})\ln\left(\frac{V_{vs}^{dyn}}{V_{vs}}\right). \quad (11)$$

(Kaneko et al., 2010).

Equation 11 helps us to understand the numerical results presented so far. The average value of V_{vs} over VS patches (shown in Figure 8) becomes larger prior to a mainshock. From Equation 11, $\Delta\tau_{vs}$ is smaller when the VS patches slip faster (i.e., larger τ_{vs} and hence smaller $\Delta\tau_{vs}$) right before the mainshock, promoting the earthquake rupture to propagate through the VS patches.

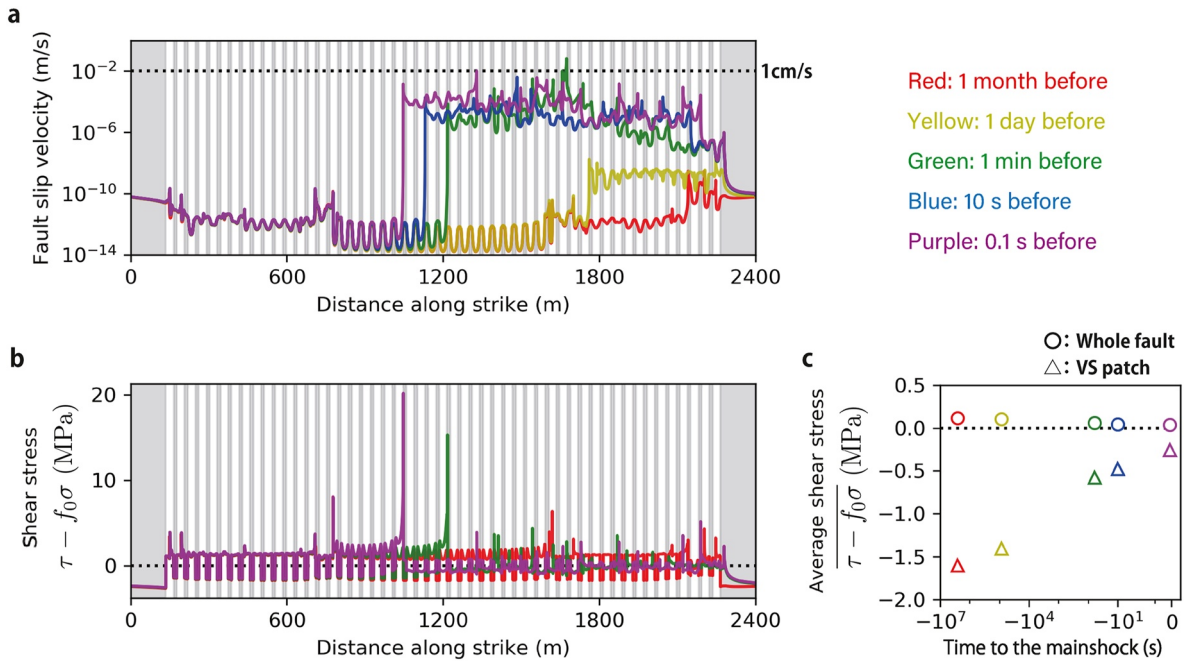

Figure 8. Evolution of slip rates and shear stress during the active foreshock period in the case shown in Figure 4. The snapshot of (a) slip rates and (b) shear stresses along the fault. Red, yellow, green, blue, and purple represent 1 month before, 1 day before, 1 min before, 10 s before, and 0.1 s before the mainshock, respectively. Gray and white areas represent velocity strengthening (VS) and velocity weakening patches, respectively. Shear stresses are shown in only three instances (red, green, and purple), for clarity. (c) Spatial average of shear stresses, $\overline{\tau - f_0\sigma}$, with time. Circles and triangles represent the spatial average over the entire fault and the only VS area, respectively. The average shear stress on the entire fault remains nearly constant over time while the average shear stress over the VS patches increases with time.

In summary, b_{GR} decreases with time for the following reasons. Due to rupturing single or a few VW patches (i.e., small foreshocks), afterslip occurs in the surrounding VS patches. These VS patches have larger slip velocities V_{vs} and larger shear stresses τ_{vs} than other VS patches do (Figure 8). Subsequent afterslip in the VS patches triggers the nucleation of a foreshock in the neighboring VW patch. Once foreshock rupture reaches the VS patches that are creeping (i.e., larger V_{vs}, larger τ_{vs}, and smaller $\Delta\tau_{vs}$), the rupture tends to propagate through the VS patches and expands spatially (Figures 9b and 9c). Repeated foreshocks make the regions of higher V_{vs} to expand spatially (Figure 8a), and consequently, the likelihood for a larger event to occur increases statistically, and b_{GR} decreases with time prior to the mainshock (Figure 6).

5. Dependence of the Temporal Change of b_{GR} on the Degree of Frictional Heterogeneities

5.1. Amplitude of b_{GR} Change Depends on the Friction Properties of VS Patches

To examine how the degree of fault frictional heterogeneities influences the resulting b_{GR} changes, we consider cases with different values of $a_{vs} - b_{vs}$ in the VS patches. In Section 3, we find that in the case where $a_{vs} - b_{vs}$ is small, fewer foreshocks are generated over mainshock cycles (Figure 2a). We first zoom in and plot the distributions of foreshocks over the final 24 hr before the mainshock in that case (Figure S3 in Supporting Information S1). Compared to the case shown in Figure 4, the case with smaller $a_{vs} - b_{vs}$ results in a few small foreshocks prior to the nucleation of a mainshock that ruptures the entire fault segment (Figure S3 in Supporting Information S1). Since only a few foreshocks occur in this case, there are too few foreshocks to analyze their statistics.

We then gradually vary $a_{vs} - b_{vs}$ from the case of accelerating foreshocks (Figure 2b) while keeping σ constant, and calculate the relative rate of b_{GR} decrease, $\Delta b_{GR}/b_{GR}$, from a reference time to the mainshock time, where Δb_{GR} is the change in b_{GR} over the specified time window. Figure 10a and Figure S4 in Supporting Information S1 show how parameter $a_{vs} - b_{vs}$ affects both the relative and absolute change in b_{GR}, respectively. When $a_{vs} - b_{vs}$ is large, the absolute value of b_{GR} increases but Δb_{GR} increases even more, and hence the relative change $\Delta b_{GR}/b_{GR}$ becomes larger (Figure 10a and Figure S4 in Supporting Information S1). Increasing $a_{vs} - b_{vs}$ further leads to swarm-like behavior as described in Section 3, making it difficult to define which events are the mainshocks.

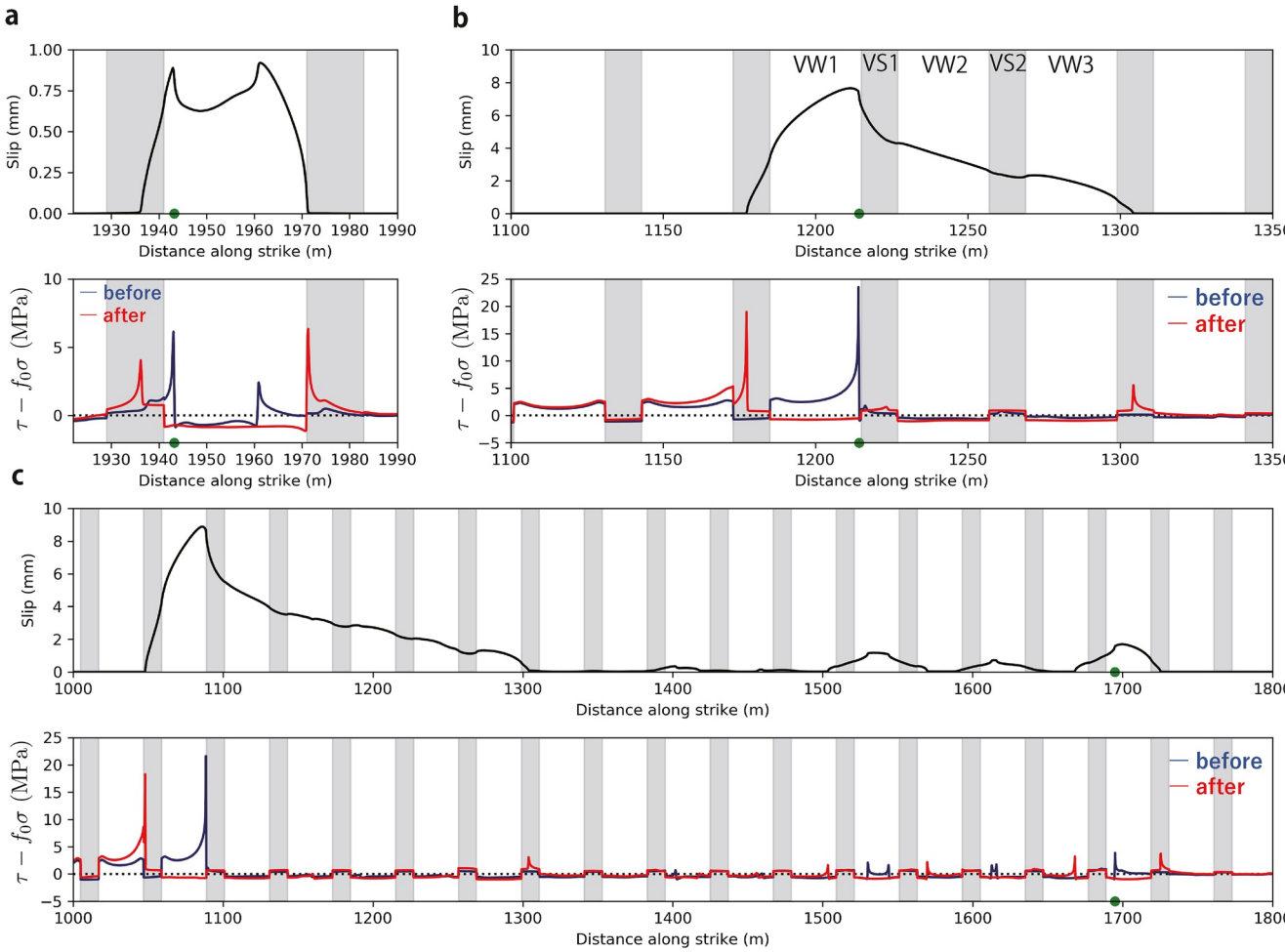

Figure 9. Coseismic slip and shear stress distributions before and after selected foreshocks. Panel (a–c) correspond to Foreshocks P, Q, and R in Figure 4, respectively. The M_w of Foreshocks P, Q, and R are 2.0, 2.9, and 3.1, respectively. Shear stress, $\tau - f_0\sigma$, before the foreshock is shown in blue and after the foreshock is shown in red. The gray and white background areas show the velocity strengthening (VS) and velocity weakening (VW) patches, respectively. Green dots represent the hypocenters of each foreshock. Foreshock rupture nucleates within a VW patch and arrests in the VS patch. VW patches VW1, VW2, and VW3, and VS patches VS1 and VS2 in panel (b) are labeled and referred to in the main text. A larger foreshock ruptures through multiple VS patches where the shear stress before the event is relatively large.

We suspect that the case marked by the dashed square in Figure 10a is out of the general trend due to the uncertainties in defining the mainshocks. Also, when $a_{vs} - b_{vs}$ becomes too small, there are not enough foreshocks to be included to calculate b_{GR}. Note that in producing a frequency-magnitude distribution (e.g., Figure 6), event stacking is performed relative to the time to the quasi-periodic mainshocks. Since much of b_{GR} changes occur over the final few percent of the mainshock recurrence intervals, the calculated b_{GR} changes are not affected by the stacking procedure. In summary, these results show that frictional heterogeneities and the rate of b_{GR} decrease are approximately positively correlated regardless of chosen time windows (Figure 10a, Figures S5a and S5d in Supporting Information S1). This means that stronger frictional heterogeneities lead to a faster b_{GR} decrease.

The time at which $\Delta b_{GR}/b_{GR}$ begins to sharply decrease also depends on the friction properties. For the cases with larger $a_{vs} - b_{vs}$, $\Delta b_{GR}/b_{GR}$ starts decreasing sooner (Figure S5 in Supporting Information S1). Hence, the duration of b_{GR} change is longer for faults that are prone to creep.

The correlation between the rate of decrease in b_{GR} and the degree of frictional heterogeneities can be explained by the amount of average shear stress increase within the VS patches (Figure 10b). The change in the average shear stress over the entire fault is almost constant, regardless of frictional heterogeneities and chosen time windows (open circles in Figure 10b). In contrast, the average shear stress over the VS patches is larger with the increasing degree of frictional heterogeneities regardless of chosen time windows (Figure 10b, Figures S5b and

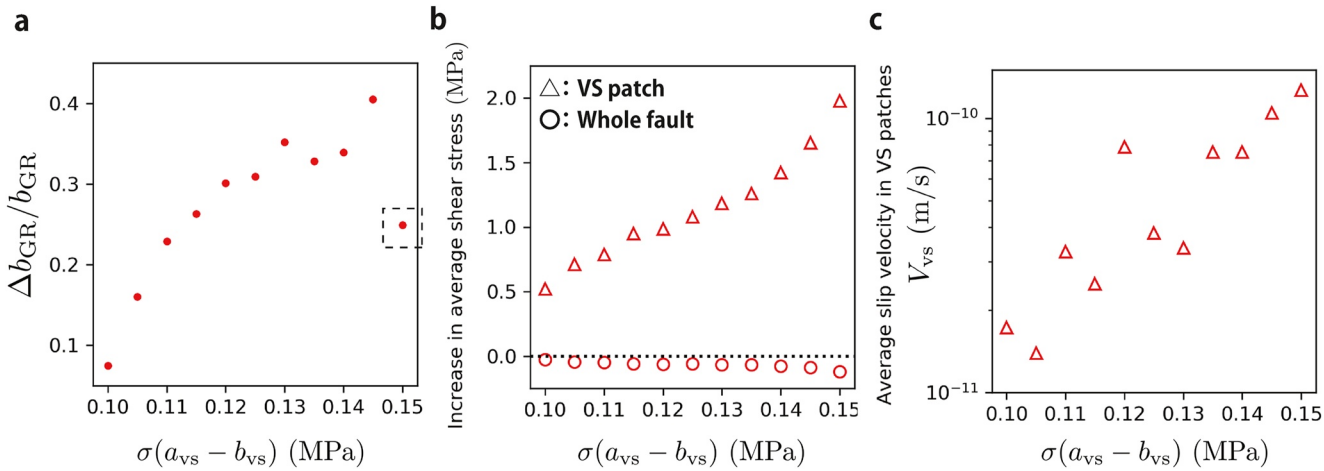

Figure 10. The dependence of b_{GR} on different distributions of $\sigma(a_{vs} - b_{vs})$. (a) The rate of b_{GR} change expressed as $\Delta b_{GR}/b_{GR}$ in the cases with a different distribution of $\sigma(a_{vs} - b_{vs})$. Each dot represents the b_{GR} change from 1 month before the mainshock (e.g., red dots in Figure 6) to 0.1 s before the mainshock (e.g., purple dots in Figure 6). The rate of decrease in b_{GR} value correlates with the increasing degree of frictional heterogeneities. (b) Amount of increase in average shear stress over the same time period as in panel (a). Circles indicate average values over the entire fault, and triangles indicate average values over the velocity strengthening (VS) patches. The amount of shear stress increase in the VS patches positively correlates with the degree of frictional heterogeneities. (c) Average slip velocity over the same time period as in panel (a). The average slip velocity in the VS patches generally increases with the degree of frictional heterogeneities.

S5e in Supporting Information S1). The average slip velocity in the VS region, V_{vs}, also increases as the degree of frictional heterogeneities increases (Figure 10c, Figures S5c and S5f in Supporting Information S1). Hence, a larger degree of frictional heterogeneities leads to a larger increase in the average shear stress in the VS patches, which in turn, promotes a larger reduction of b_{GR}.

5.2. Amplitude of b_{GR} Change Also Depends on the Friction Properties of VW Patches

We also investigate how frictional heterogeneities in the VW patches $a_{vw} - b_{vw}$ affect the change in b_{GR}. The case with stronger velocity-weakening patches (i.e., more negative $a_{vw} - b_{vw}$) while keeping the same VS properties leads to smaller $\Delta b_{GR}/b_{GR}$ (Figure 11a and Figure S6 in Supporting Information S1). As the effectiveness of a rupture barrier also depends on the degree of VW (Kaneko et al., 2010), increasing the degree of VW induces a similar effect as decreasing $a_{vs} - b_{vs}$ in the VS patches (Figure 11a). The tendency for earthquake rupture to propagate through a VS patch increases with more negative $a_{vw} - b_{vw}$, resulting in fewer foreshocks. Figure 11b shows that changing VW properties leads to increased average shear stress over the VS patches, with a positive correlation between the relative change of b_{GR} and the average shear stress over VS patches in these cases. Hence,

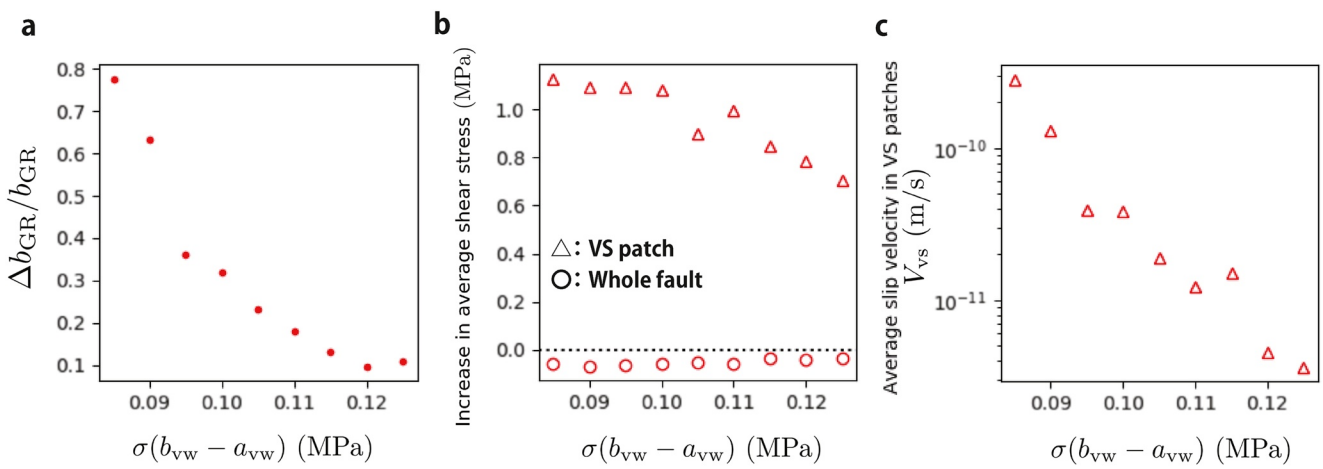

Figure 11. The dependence of b_{GR} on different distributions of $\sigma(b_{vw} - a_{vw})$. (a) The rate of b_{GR} change expressed as $\Delta b_{GR}/b_{GR}$ in the cases with a different distribution of $\sigma(b_{vw} - a_{vw})$. (b) Amount of increase in average shear stress over the same time period as in panel (a). (c) Average slip velocity over the same time period as in panel (a).

the temporal change of b_{GR} depends not only on the properties of creeping patches but also on the properties of locked (VW) fault patches.

5.3. Effect of the Absolute Level of Shear Strength on b_{GR} Change

In the models presented so far, the reference friction coefficient $f_0 = 0.6$ and is uniform over the fault. Rock-friction laboratory experiments show that the absolute value of friction f_0 depends on the friction properties (Ikari, Marone, & Saffer, 2011). f_0 can be smaller in VS materials while f_0 in VW materials remains high ($f_0 \sim 0.6$). This means that VS fault patches may be weaker than VW patches. To test how the absolute strength of a fault influences the temporal changes in b_{GR}, we consider an additional model where $f_0 = 0.3$ in the VS patches and $f_0 = 0.6$ in the VW patches (Figure 12). We find that the resulting foreshock behavior and b_{GR} remain unchanged with respect to the case with uniform $f_0 = 0.6$ (Compare Figures 12e and 6). This is somewhat expected, as models of earthquake cycles based on rate-and-state friction do not depend on the values of reference friction coefficient f_0 and hence the absolute strength. This suggests that our proposed mechanism of b_{GR} changes may be applicable to both weak and strong faults.

6. Discussion

6.1. On the Scale Dependence of Spatio-Temporal Evolution of Foreshocks

To gain insights into what controls the time and spatial scales of foreshocks, we vary the value of characteristic slip D_{RS}. In solving equations of dynamic elasticity coupled with a rate-and-state fault, one can non-dimensionalize time and space via $t \rightarrow t' = tV_0/D_{RS}$ and $x \rightarrow x' = x/D_{RS}$, where t' and x' are non-dimensional time and space, respectively. Then the results expressed with t' and x' do not change (e.g., Tape et al., 2018) such that the dimensional time and space in the corresponding calculation results scale with D_{RS} without conducting simulations with large or small D_{RS} (e.g., Tape et al., 2018). Figure 13 shows that the models with significantly small or large D_{RS} qualitatively reproduce the spatial and temporal scales of foreshocks reported in nature and laboratory experiments despite several orders of magnitude differences in the time and spatial scales between the Tohoku-Oki earthquake and laboratory experiments. This means that the orders-of-magnitude differences in the spatial and temporal scales of foreshocks in nature and laboratory experiments might simply reflect different D_{RS} on natural and laboratory faults. Since D_{RS} may depend on the microscopic asperity size of a fault (e.g., Marone & Saffer, 2015), our results suggest that the duration of b_{GR} changes observed on natural faults (Table 1) may scale with the microscopic asperity size of a fault.

Note that, to better match the amplitude of a b_{GR} change and the time scale for a specific observation (e.g., Figure 13), other parameters such as the background loading rate and friction parameters need to be adjusted. For example, the time scale is inversely proportional to the background loading rate (Kaneko et al., 2016, 2017), and our assumed tectonic loading rate is too slow to correctly reproduce the time scale of the laboratory experiment (Figures 13d and 13e). In addition, the 2D model and regular distribution of frictional heterogeneities assumed here are oversimplified, and a more realistic model would be needed to accurately reproduce a specific observation or laboratory experiment.

6.2. Frictional Heterogeneities Versus Normal Stress Heterogeneties

In the previous numerical study of earthquake cycles on a rough fault, Cattania and Segall (2021) showed that heterogeneous normal stresses induced by fault surface roughness can result in active and accelerating foreshocks. While we have not yet considered normal stress heterogeneities on a planar fault, the key phenomenon that controls the temporal changes in b_{GR} is the occurrence of aseismic slip. In our model, VS patches induce aseismic slip, which derives active and accelerating foreshocks. Even on a fault with uniform friction properties (i.e., uniform $a_{vw} - b_{vw}$), if the normal stress is heterogeneous such that the nucleation size is locally much larger, resulting aseismic slip could inhibit earthquake rupture propagation, leading to accelerating foreshocks and a temporal change in b_{GR}. Hence any heterogeneities in parameters that control the nucleation size (such as μ, σ, etc) have the potential to produce accelerating foreshock behavior and temporal changes in b_{GR}.

6.3. On Insufficient Number of Events to Robustly Estimate b_{GR}

While parameter b_{GR} is widely used to monitor the change in seismic activity, it is important to include many (e.g., hundreds) of foreshocks in robustly estimating b_{GR} changes, which is often challenging for real earthquake

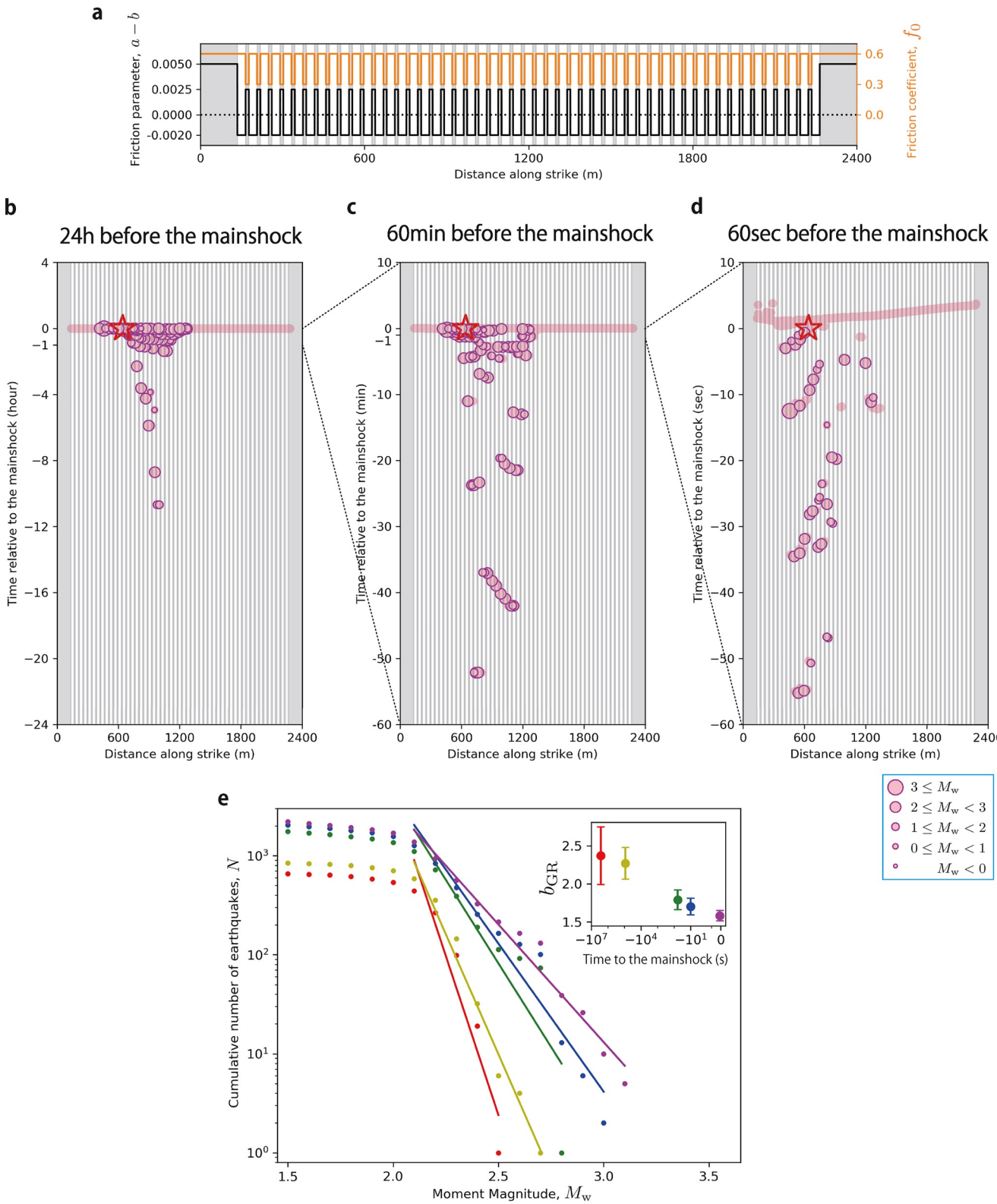

Figure 12. Spatio-temporal evolution of foreshocks and temporal change of b_{GR} in the case with non-uniform reference friction coefficient f_0. (a) The distributions of rate and state parameters $a - b$ and f_0. In this case, $f_0 = 0.3$ in the velocity strengthening patches, and $f_0 = 0.6$ in the velocity weakening patches, as motivated by laboratory results of Ikari, Marone, and Saffer (2011). Foreshock distributions over the time periods of (b) 24 hr, (c) 60 min, and (d) 60 s before the mainshock. (e) Relationship between M_w and cumulative number of earthquakes N plotted over five-time windows. The inset figure shows the time variation of b_{GR} obtained by the maximum likelihood estimation (Equation 8) with $M_c = 2.1$. The values of f_0 do not affect the resulting foreshock behavior and b_{GR}.

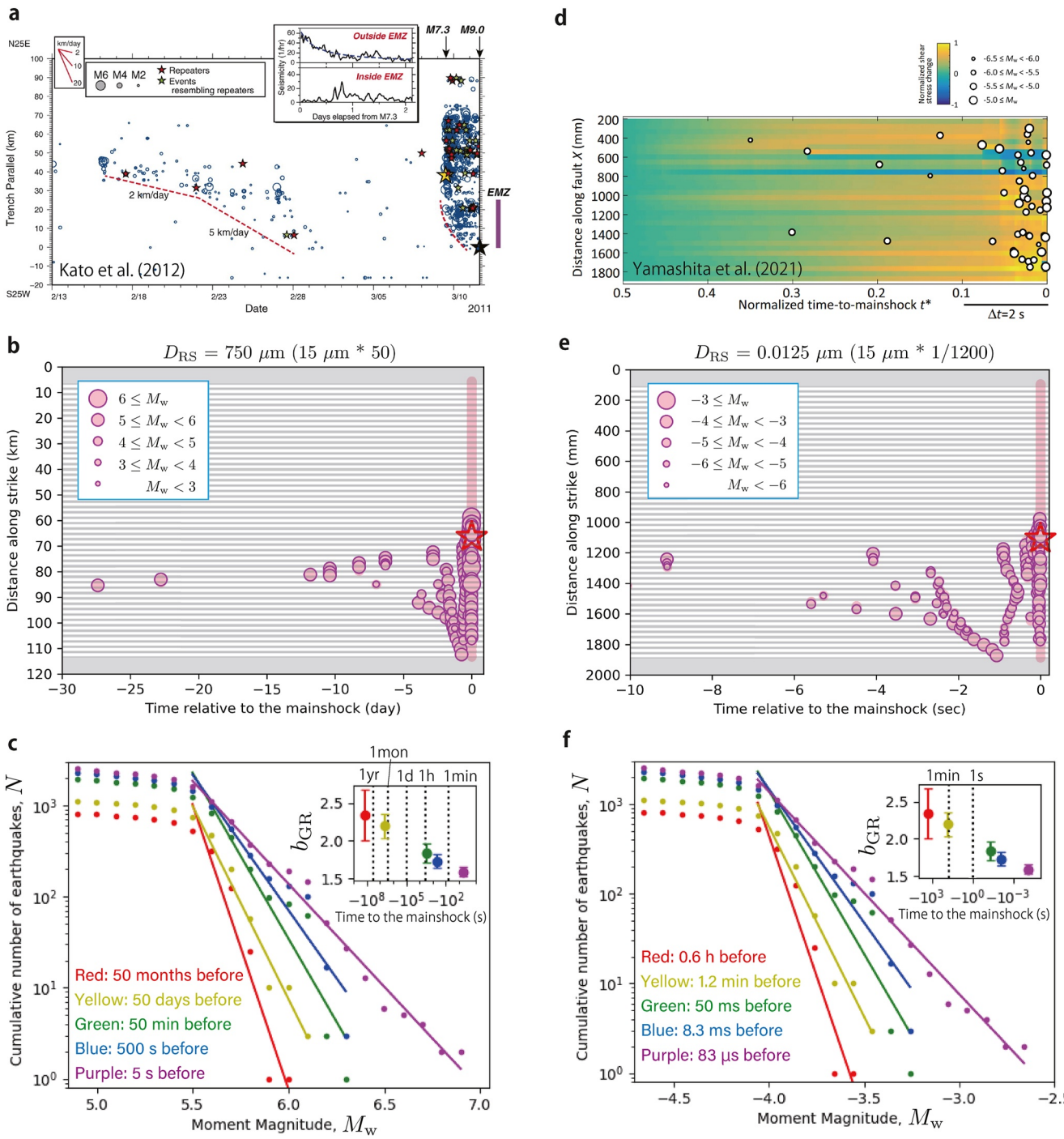

Figure 13. Comparisons of the spatio-temporal evolution of foreshocks between (a–c) nature and model with D_{RS} = 750 μm and (d–f) laboratory experiment and model with D_{RS} = 0.0125 μm. The case with D_{RS} = 750 μm, which is 50 times larger than the case with D_{RS} = 15 μm in Figure 4, is compared to the distribution of foreshocks for the 2011 M_w9.0 Tohoku-Oki earthquake (Kato et al., 2012). The case with D_{RS} = 0.0125 μm, which is 1,200 times smaller than the case shown in Figure 4, is compared to the distribution of foreshocks (acoustic emission events, open circles) in the laboratory experiment of Yamashita et al. (2021). The spatial and temporal scales of foreshock evolution in these models are qualitatively similar to those in nature and laboratory experiments.

catalogs. The cumulative number of foreshocks in a given time window has been commonly used to generate the magnitude-frequency distributions and corresponding b_{GR} estimates (e.g., Nanjo et al., 2012). We find that using the cumulative number of foreshocks since a reference time yields more robust estimates of b_{GR} changes as opposed to the cumulative number of foreshocks in a given time window (Compare Figure 6 and Figure S2 in

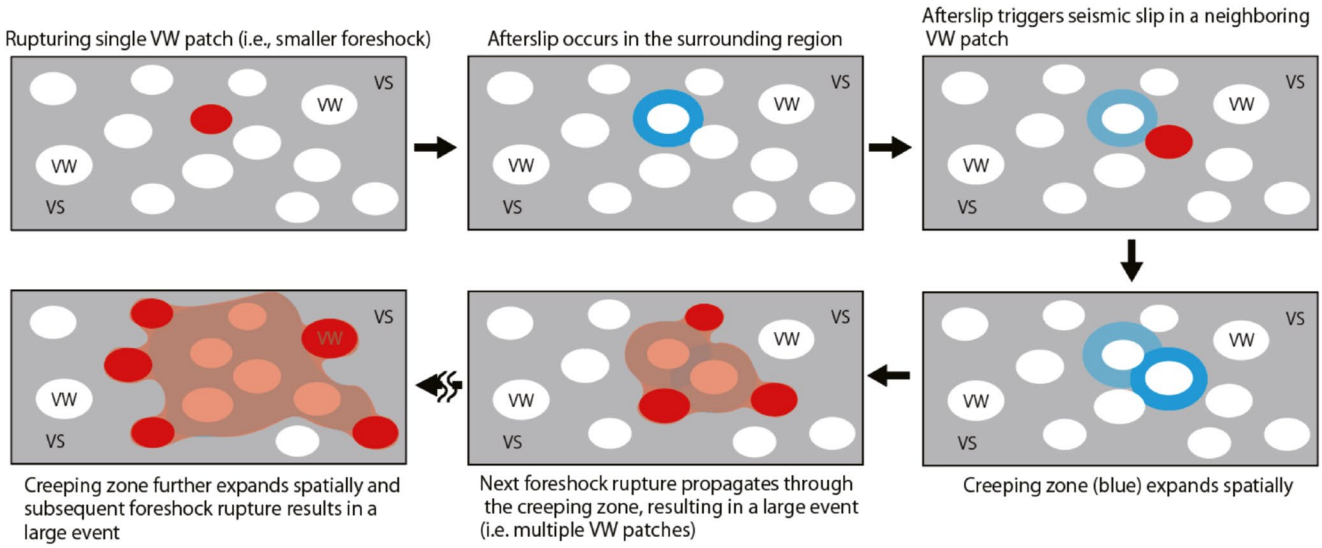

Figure 14. Cartoon illustrating how b_{GR} decreases with time toward the mainshock. Background gray and white areas represent velocity strengthening (VS) and velocity weakening (VW) fault patches, respectively. Red represents coseismic slip regions, and blue indicates afterslip regions. Orange also represents coseismic slip regions which include both VS patches that were slipping at a faster slip rate and VW patches that were already ruptured during previous foreshocks. The larger the size of the combined red and orange regions is, the larger the foreshock. As the region of afterslip due to foreshocks expands, a foreshock size becomes larger on average, leading to the decrease of b_{GR} and increasing the likelihood of future rupture to propagate through the foreshock zone (i.e., becoming the mainshock).

Supporting Information S1). From the viewpoint of securing a sufficient number of events, such an approach may be useful in robustly estimating the temporal changes of b_{GR} in nature and laboratory experiments.

6.4. Model Limitation

We have made several important simplifications in our model. First, the model only includes a single planar fault embedded into the Earth's crust, producing various sizes of earthquakes. In contrast, b_{GR} changes inferred from observational studies contain seismicity on subsidiary faults and faults in damaged zones in the mainshock source region. Incorporating subsidiary faults (e.g., Ozawa & Ando, 2021) into models of earthquake cycles and studying their influence on b_{GR} changes remain a subject of future studies. Second, unlike natural faults, our model produces few or no aftershocks (e.g., Figure 4). This problem could be solved if the model includes subsidiary faults (e.g., Ozawa & Ando, 2021). Third, we consider a regular distribution of frictional heterogeneities, for simplicity. On natural faults, VS and VW patches are likely distributed in a fractal manner, as considered in models of earthquake cycles (e.g., Aochi & Ide, 2009; Dublanchet, 2022). While assuming a fractal distribution of frictional heterogeneities in the fully dynamic model of earthquake cycles and comparing their resulting b-value evolution to that with the simplified heterogeneities is a subject of future work, our model result demonstrates that a power law distribution of foreshocks can emerge without the need for a fractal distribution of frictional heterogeneities (Figure 6). Finally, we assume that the fault constitutive response is represented by the standard rate-and-state friction laws without accounting for the enhanced weakening at seismic slip rates (~1 m/s) (e.g., Di Toro et al., 2011; Noda & Lapusta, 2013; Rice, 2006). How and to what extent enhanced coseismic weakening affects the temporal changes in b_{GR} remains unclear and a subject of future work.

6.5. Proposed Mechanism of the Temporal Decrase of b_{GR}

From the results and discussion thus far, we summarize the underlying process of b_{GR} changes that may explain the widely reported observations of the temporal decrease of b_{GR} as follows: (a) Due to rupturing a single or few VW patches (i.e., smaller foreshock), afterslip occurs in the surrounding VS region. That region has larger slip velocities V_{vs} and larger shear stresses τ_{vs} than the other VS patches (Figure 14). (b) The afterslip of the previous foreshock triggers seismic rupture in the neighboring VW patch, and the afterslip regions (with larger V_{vs}, larger τ_{vs}, and smaller $\Delta\tau_{vs}$) expand spatially, which increases the connectivity of VW asperities via seismic rupture (Figure 14). (c) When the next foreshock occurs, the rupture can propagate through the aseismically slipping

region, resulting in a larger foreshock (i.e., multiple VW patches) (Figure 14). (d) Afterslip of that foreshock and other foreshocks will further expand, resulting in a higher probability of a larger seismic event (Figure 14). As the foreshock zone acts as a less effective barrier against future coseismic rupture, the mainshock rupture with the nucleation indistinguishable from other foreshocks tends to propagate through the foreshock zone, and hence the mainshock area becomes larger than the foreshock zone. This mechanism may explain how and why mainshock rupture becomes large, instead of resulting in another foreshock (i.e., a smaller event).

6.6. Implications for b_{GR} Changes Reported in Laboratory Studies, Other Earthquake Phenomena and Short-Term Earthquake Forecasting

The proposed mechanism of a b_{GR} decrease is consistent with the laboratory results of Bolton et al. (2021) in that a decrease of b_{GR} correlates with increasing fault slip velocity (V_{vs} in our model). In addition, the temporal decrease of b_{GR} reported in the laboratory experiments of Yamashita et al. (2021) can be interpreted based on the proposed mechanism (Figure 14). Our results suggest that undetected, minor aseismic slip induced by heterogeneously distributed fault gouge (Yamashita et al., 2021) might have been the reason why the b_{GR} decrease was observed prior to the mainshock when the nominal shear stress on the fault was nearly constant.

Our modeling results can also be used to interpret seemingly unrelated observations of the change in the source characteristics of repeating earthquakes. Uchida et al. (2015) inferred the enlargement of the size of repeating earthquakes following the Tohoku-Oki earthquake due to increased loading in the surrounding creeping zone. Assuming that these repeating earthquakes occur on an isolated VW patch surrounded by the stable VS fault segment that acts as a barrier to the seismic rupture, our analytical result presented in Section 4 suggests that faster creep rates in the VS regions due to the increased loading can reduce the rupture barrier effectiveness and promote the propagation of seismic rupture in these regions, leading to a larger size of repeating earthquakes. Hence the proposed mechanism of temporal change in b_{GR} may be applicable to other related earthquake phenomena.

In Section 5, we find that the absolute value of b_{GR} increases with increasing $a_{vs} - b_{vs}$ on the VS fault patches. Since larger $a_{vs} - b_{vs}$ leads to faster fault slip rates (Figure 10), the model behavior is qualitatively consistent with the observation of Vorobieva et al. (2016) in that b_{GR} is larger on fault segments with faster creep rates. Our results further suggest that temporal changes of b_{GR} are likely more often observed on "mildly" creeping faults. In order for accelerating foreshocks to occur, creeping patches (i.e., VS patches in our model) need to be in the transitional regime in between fewer foreshocks and earthquake swarm behaviors (Figure 2) and act as mild barriers to earthquake ruptures. Within this transitional regime, a fault that is prone to more creep, leads to a larger change of b_{GR} (Figure 10). This may explain why large earthquakes in nature often accompany few or no foreshocks and that those that are prone to creep (e.g., plate-boundary faults) tend to exhibit the temporal change of b_{GR} before the mainshock. This inference is supported by the observations that both the Tohoku-Oki and Iquique foreshocks and the resulting b_{GR} decrease occurred at plate boundaries where slow earthquakes have also been reported (e.g., Y. Ito et al., 2013; Schurr et al., 2014). Lastly, our results that a faster rate of b_{GR} decrease occurs during the time closer to the mainshock (Figure 7) implies that the detailed monitoring of b_{GR} may be used to accurately forecast the timing of a future mainshock.

7. Conclusions

Models with the heterogeneous distribution of friction properties lead to various earthquake phenomena (i.e., earthquakes with few foreshocks, earthquakes with active foreshocks, earthquake swarms, etc). We have identified a parameter regime in which the model gives rise to an accelerating foreshock behavior prior to a mainshock and analyzed the foreshock statistics. We have found that, like in observations, a complex pattern of migrating foreshocks can be seen in this model and that b_{GR} decreases with time prior to the mainshock. In this case, increased shear stresses within the creeping (or velocity-strengthening) patches due to numerous foreshocks make these patches more susceptible to future co-seismic slip, increasing the likelihood of large ruptures and resulting in smaller b_{GR} with time. In other words, the temporal decrease of the b-value is caused by the reduction of rupture barrier effectiveness due to increasing shear stresses within slowly creeping regions. At the same time, our model shows that the average shear stresses over the entire fault are nearly constant with time during the accelerating foreshocks. Our results further suggest that the overall asperity size of a fault may control the duration of b_{GR} changes. We conclude that a potential physical explanation for the temporal decrease of b_{GR} may be an increased level of shear stresses in aseismically slipping regions that make these regions more susceptible to co-seismic slip, leading to larger foreshocks.

Data Availability Statement

The numerical data used to produce all the figures are available at R. Ito and Kaneko (2023).

Acknowledgments
We thank Keiko Kuge, Bogdan Enescu, Hiroyuki Noda, and Reiju Norisugi for their helpful discussions. We also thank Editor Rachel Abercrombie, the Associate Editor, and two anonymous reviewers for their comments that helped us greatly improve the manuscript. This work was supported by the JSPS KAKENHI (21H05206 and 23K03547).

References

Aki, K. (1965). Maximum likelihood estimate of *b* in the formula log $N = a - bM$ and its confidence limits. *Bulletin of the Earthquake Research Institute, University of Tokyo*, *43*, 237–239.

Ando, R., & Imanishi, K. (2011). Possibility of Mw 9.0 mainshock triggered by diffusional propagation of after-slip from Mw 7.3 foreshock. *Earth Planets and Space*, *63*(7), 767–771. https://doi.org/10.5047/eps.2011.05.016

Aochi, H., & Ide, S. (2009). Complexity in earthquake sequences controlled by multiscale heterogeneity in fault fracture energy. *Journal of Geophysical Research*, *114*(B3), B03305. https://doi.org/10.1029/2008jb006034

Beall, A., van den Ende, M., Ampuero, J.-P., Capitanio, F. A., & Fagereng, A. (2022). Linking earthquake magnitude-frequency statistics and stress in visco-frictional fault zone models. *Geophysical Research Letters*, *49*(20), e2022GL099247. https://doi.org/10.1029/2022gl099247

Bolton, D. C., Shreedharan, S., Rivière, J., & Marone, C. (2021). Frequency-magnitude statistics of laboratory foreshocks vary with shear velocity, fault slip rate, and shear stress. *Journal of Geophysical Research: Solid Earth*, *126*(11), e2021JB022175. https://doi.org/10.1029/2021JB022175

Bouchon, M., Karabulut, H., Aktar, M., Özalaybey, S., Schmittbuhl, J., & Bouin, M.-P. (2011). Extended nucleation of the 1999 M_w 7.6 Izmit earthquake. *Science*, *331*(6019), 877–880. https://doi.org/10.1126/science.1197341

Brodsky, E. E., & Lay, T. (2014). Recognizing foreshocks from the 1 April 2014 Chile earthquake. *Science*, *344*(6185), 700–702. https://doi.org/10.1126/science.1255202

Cattania, C., & Segall, P. (2021). Precursory slow slip and foreshocks on rough faults. *Journal of Geophysical Research: Solid Earth*, *126*(4), e2020JB020430. https://doi.org/10.1029/2020jb020430

Dascher-Cousineau, K., Lay, T., & Brodsky, E. E. (2020). Two foreshock sequences Post Gulia and Wiemer (2019). *Seismological Research Letters*, *91*(5), 2843–2850. https://doi.org/10.1785/0220200082

Day, S. M., Dalguer, L. A., Lapusta, N., & Liu, Y. (2005). Comparison of finite difference and boundary integral solutions to three-dimensional spontaneous rupture. *Journal of Geophysical Research*, *110*(B12), B12307. https://doi.org/10.1029/2005JB003813

Dieterich, J. H. (1978). Time-dependent friction and the mechanics of stick-slip. *Pure and Applied Geophysics*, *116*(4–5), 790–806. https://doi.org/10.1007/BF00876539

Dieterich, J. H. (1979). Modeling of rock friction: 1. Experimental results and constitutive equations. *Journal of Geophysical Research*, *84*(B5), 2161–2168. https://doi.org/10.1029/JB084iB05p02161

Dieterich, J. H. (1992). Earthquake nucleation on faults with rate- and state-dependent strength. *Tectonophysics*, *211*(1–4), 115–134. https://doi.org/10.1016/0040-1951(92)90055-B

Di Toro, G., Han, R., Hirose, T., Paola, N. D., Nielsen, S., Mizoguchi, K., et al. (2011). Fault lubrication during earthquakes. *Nature*, *471*(7339), 494–498. https://doi.org/10.1038/nature09838

Dodge, D. A., Beroza, G. C., & Ellsworth, W. L. (1996). Detailed observations of California foreshock sequences: Implications for the earthquake initiation process. *Journal of Geophysical Research*, *101*(B10), 22371–22392. https://doi.org/10.1029/96jb02269

Dublanchet, P. (2018). The dynamics of earthquake precursors controlled by effective friction. *Geophysical Journal International*, *212*(2), 853–871. https://doi.org/10.1093/gji/ggx438

Dublanchet, P. (2020). Stress-dependent *b* value variations in a heterogeneous rate-and-state fault model. *Geophysical Research Letters*, *47*(13), e2020GL087434. https://doi.org/10.1029/2020GL087434

Dublanchet, P. (2022). Shear stress and *b*-value fluctuations in a hierarchical rate-and-state asperity model. *Pure and Applied Geophysics*, *179*(6–7), 2423–2435. https://doi.org/10.1007/s00024-022-03039-3

Dublanchet, P., Bernard, P., & Favreau, P. (2013). Interactions and triggering in a 3-D rate-and-state asperity model. *Journal of Geophysical Research: Solid Earth*, *118*(5), 2225–2245. https://doi.org/10.1002/jgrb.50187

Ellsworth, W. L., & Bulut, F. (2018). Nucleation of the 1999 Izmit earthquake by a triggered cascade of foreshocks. *Nature Geoscience*, *11*(7), 531–535. https://doi.org/10.1038/s41561-018-0145-1

Goebel, T. H. W., Schorlemmer, D., Becker, T. W., Dresen, G., & Sammis, C. G. (2013). Acoustic emissions document stress changes over many seismic cycles in stick-slip experiments. *Geophysical Research Letters*, *40*(10), 2049–2054. https://doi.org/10.1002/grl.50507

Gulia, L., Rinaldi, A. P., Tormann, T., Vannucci, G., Enescu, B., & Wiemer, S. (2018). The effect of a mainshock on the size distribution of the aftershocks. *Geophysical Research Letters*, *45*(24), 13277–13287. https://doi.org/10.1029/2018GL080619

Gulia, L., Tormann, T., Wiemer, S., Herrmann, M., & Seif, S. (2016). Short-term probabilistic earthquake risk assessment considering time-dependent *b* values. *Geophysical Research Letters*, *43*(3), 1100–1108. https://doi.org/10.1002/2015GL066686

Gulia, L., & Wiemer, S. (2019). Real-time discrimination of earthquake foreshocks and aftershocks. *Nature*, *574*(7777), 193–199. https://doi.org/10.1038/s41586-019-1606-4

Gutenberg, B., & Richter, C. (1949). *Seismicity of the earth and associated phenomena*. Princeton University Press.

Harris, R. A. (2017). Large earthquakes and creeping faults. *Reviews of Geophysics*, *55*(1), 169–198. https://doi.org/10.1002/2016rg000539

Ikari, M. J., Marone, C., & Saffer, D. M. (2011). On the relation between fault strength and frictional stability. *Geology*, *39*(1), 83–86. https://doi.org/10.1130/G31416.1

Ikari, M. J., Niemeijer, A. R., & Marone, C. (2011). The role of fault zone fabric and lithification state on frictional strength, constitutive behavior, and deformation microstructure. *Journal of Geophysical Research*, *116*(B8), B08404. https://doi.org/10.1029/2011jb008264

Ito, R., & Kaneko, Y. (2023). Numerical data for research paper titled "Physical mechanism for a temporal decrease of the Gutenberg-Richter *b*-value prior to a large earthquake" (version 2) [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.8115932

Ito, Y., Hino, R., Kido, M., Fujimoto, H., Osada, Y., Inazu, D., et al. (2013). Episodic slow slip events in the Japan subduction zone before the 2011 Tohoku-Oki earthquake. *Tectonophysics*, *600*, 14–26. https://doi.org/10.1016/j.tecto.2012.08.022

Kanamori, H. (1977). The energy release in great earthquakes. *Journal of Geophysical Research*, *82*(20), 2981–2987. https://doi.org/10.1029/jb082i020p02981

Kaneko, Y., Avouac, J.-P., & Lapusta, N. (2010). Towards inferring earthquake patterns from geodetic observations of interseismic coupling. *Nature Geoscience*, *3*(5), 363–369. https://doi.org/10.1038/NGEO843

Kaneko, Y., Carpenter, B. M., & Nielsen, S. B. (2017). Nucleation process of magnitude 2 repeating earthquakes on the San Andreas Fault predicted by rate-and-state fault models with SAFOD drill core data. *Geophysical Research Letters*, *44*(1), 162–173. https://doi.org/10.1002/2016GL071569

Kaneko, Y., & Lapusta, N. (2008). Variability of earthquake nucleation in continuum models of rate-and-state faults and implications for aftershock rates. *Journal of Geophysical Research*, *113*(B12), B12312. https://doi.org/10.1029/2007JB005154

Kaneko, Y., Lapusta, N., & Ampuero, J.-P. (2008). Spectral element modeling of spontaneous earthquake rupture on rate and state faults: Effect of velocity-strengthening friction at shallow depths. *Journal of Geophysical Research*, *113*(B9), B09317. https://doi.org/10.1029/2007JB005553

Kaneko, Y., Nielsen, S. B., & Carpenter, B. M. (2016). The onset of laboratory earthquakes explained by nucleating rupture on a rate-and-state fault. *Journal of Geophysical Research: Solid Earth*, *121*(8), 6071–6091. https://doi.org/10.1002/2016JB013143

Kato, A., Obara, K., Igarashi, T., Tsuruoka, H., Nakagawa, S., & Hirata, N. (2012). Propagation of slow slip leading up to the 2011 M_w 9.0 Tohoku-Oki earthquake. *Science*, *335*(6069), 705–708. https://doi.org/10.1126/science.1215141

Lapusta, N., & Liu, Y. (2009). Three-dimensional boundary integral modeling of spontaneous earthquake sequences and aseismic slip. *Journal of Geophysical Research*, *114*(B9), B09303. https://doi.org/10.1029/2008JB005934

Lapusta, N., Rice, J. R., Ben-Zion, Y., & Zheng, G. (2000). Elastodynamic analysis for slow tectonic loading with spontaneous rupture episodes on faults with rate- and state-dependent friction. *Journal of Geophysical Research*, *105*(B10), 23765–23789. https://doi.org/10.1029/2000JB900250

Lui, S. K., & Lapusta, N. (2016). Repeating microearthquake sequences interact predominantly through postseismic slip. *Nature Communications*, *7*(1), 1–7. https://doi.org/10.1038/ncomms13020

Luo, Y., & Ampuero, J.-P. (2018). Stability of faults with heterogeneous friction properties and effective normal stress. *Tectonophysics*, *733*, 257–272. https://doi.org/10.1016/j.tecto.2017.11.006

Marone, C. (1998). Laboratory-derived friction laws and their application to seismic faulting. *Annual Review of Earth and Planetary Sciences*, *26*(1), 643–696. https://doi.org/10.1146/annurev.earth.26.1.643

Marone, C., & Saffer, D. M. (2015). The mechanics of frictional healing and slip instability during the seismic cycle. *Treatise on Geophysics*, 111–138. https://doi.org/10.1016/B978-0-444-53802-4.00092-0

Marty, S., Schubnel, A., Bhat, H. S., Aubry, J., Fukuyama, E., Latour, S., et al. (2023). Nucleation of laboratory earthquakes: Quantitative analysis and scalings. *Journal of Geophysical Research: Solid Earth*, *128*(3), e2022JB026294. https://doi.org/10.1029/2022jb026294

McLaskey, G. C. (2019). Earthquake initiation from laboratory observations and implications for foreshocks. *Journal of Geophysical Research: Solid Earth*, *124*(12), 12882–12904. https://doi.org/10.1029/2019JB018363

McLaskey, G. C., & Lockner, D. A. (2014). Preslip and cascade processes initiating laboratory stick slip. *Journal of Geophysical Research: Solid Earth*, *119*(8), 6323–6336. https://doi.org/10.1002/2014JB011220

Nanjo, K. Z., Hirata, N., Obara, K., & Kasahara, K. (2012). Decade-scale decrease in *b* value prior to the M9-class 2011 Tohoku and 2004 Sumatra quakes. *Geophysical Research Letters*, *39*(20), L20304. https://doi.org/10.1029/2012GL052997

Noda, H., & Lapusta, N. (2013). Stable creeping fault segments can become destructive as a result of dynamic weakening. *Nature*, *493*(7433), 518–521. https://doi.org/10.1038/nature11703

Ozawa, S., & Ando, R. (2021). Mainshock and aftershock sequence simulation in geometrically complex fault zones. *Journal of Geophysical Research: Solid Earth*, *126*(2), e2020JB020865. https://doi.org/10.1029/2020JB020865

Papadopoulos, G. A., Charalampakis, M., Fokaefs, A., & Minadakis, G. (2010). Strong foreshock signal preceding the L'Aquila (Italy) earthquake (M_w 6.3) of 6 April 2009. *Natural Hazards and Earth System Sciences*, *10*(1), 19–24. https://doi.org/10.5194/nhess-10-19-2010

Rice, J. R. (1993). Spatio-temporal complexity of slip on a fault. *Journal of Geophysical Research*, *98*(B6), 9885–9907. https://doi.org/10.1029/93JB00191

Rice, J. R. (2006). Heating and weakening of faults during earthquake slip. *Journal of Geophysical Research*, *111*(B5), B05311. https://doi.org/10.1029/2005JB004006

Rubin, A. M., & Ampuero, J.-P. (2005). Earthquake nucleation on (aging) rate and state faults. *Journal of Geophysical Research*, *110*(B11), B11312. https://doi.org/10.1029/2005JB003686

Ruina, A. L. (1983). Slip instability and state variable friction laws. *Journal of Geophysical Research*, *88*, 10359–10370. https://doi.org/10.1029/JB088iB12p10359

Schaal, N., & Lapusta, N. (2019). Microseismicity on patches of higher compression during larger-scale earthquake nucleation in a rate-and-state fault model. *Journal of Geophysical Research: Solid Earth*, *124*(2), 1962–1990. https://doi.org/10.1029/2018JB016395

Scholz, C. H. (1968). The frequency-magnitude relation of microfracturing in rock and its relation to earthquakes. *Bulletin of the Seismological Society of America*, *58*(1), 399–415. https://doi.org/10.1785/BSSA0580010399

Scholz, C. H. (2015). On the stress dependence of the earthquake *b* value. *Geophysical Research Letters*, *42*(5), 1399–1402. https://doi.org/10.1002/2014GL062863

Schurr, B., Asch, G., Hainzl, S., Bedford, J., Hoechner, A., Palo, M., et al. (2014). Gradual unlocking of plate boundary controlled initiation of the 2014 Iquique earthquake. *Nature*, *512*(7514), 299–302. https://doi.org/10.1038/nature13681

Schwartz, S. Y., & Rokosky, J. M. (2007). Slow slip events and seismic tremor at circum-Pacific subduction zones. *Reviews of Geophysics*, *45*(3), RG3004. https://doi.org/10.1029/2006RG000208

Shi, Y., & Bolt, B. A. (1982). The standard error of the magnitude-frequency *b* value. *Bulletin of the Seismological Society of America*, *72*(5), 1677–1687. https://doi.org/10.1785/BSSA0720051677

Shimbaru, T., & Yoshida, A. (2021). The *b* value in the seismic activity around foci of large crustal earthquakes before and after their occurrence. *Zisin 2nd Collection*, *74*, 77–86. https://doi.org/10.4294/zisin.2020-5

Simon, V., Kraft, T., Diehl, T., & Tormann, T. (2021). Possible precursory slow-slip to two ML 3 mainevents of the Diemtigen microearthquake sequence, Switzerland. *Geophysical Research Letters*, *48*(19), e2021GL093783. https://doi.org/10.1029/2021gl093783

Skarbek, R. M., Rempel, A. W., & Schmidt, D. A. (2012). Geologic heterogeneity can produce aseismic slip transients. *Geophysical Research Letters*, *39*(21), L21306. https://doi.org/10.1029/2012GL053762

Tape, C., Holtkamp, S., Silwal, V., Hawthorne, J., Kaneko, Y., Ampuero, J.-P., et al. (2018). Earthquake nucleation and fault slip complexity in the lower crust of central Alaska. *Nature Geoscience*, *1*(7), 536–541. https://doi.org/10.1038/s41561-018-0144-2

Uchida, N., Shimamura, K., Matsuzawa, T., & Okada, T. (2015). Postseismic response of repeating earthquakes around the 2011 Tohoku-Oki earthquake: Moment increases due to the fast loading rate. *Journal of Geophysical Research: Solid Earth*, *120*(1), 259–274. https://doi.org/10.1002/2013JB010933

Utsu, T. (1965). A method for determining the value of b in a formula log n=a-bM showing the magnitude frequency relation for earthquakes. *Journal of Physics of the Earth*, *14*(2), 37–40. https://doi.org/10.4294/jpe1952.14.37

Vorobieva, I., Shebalin, P., & Narteau, C. (2016). Break of slope in earthquake size distribution and creep rate along the San Andreas Fault system. *Geophysical Research Letters*, *43*(13), 6869–6875. https://doi.org/10.1002/2016gl069636

Yabe, S., & Ide, S. (2017). Slip-behavior transitions of a heterogeneous linear fault. *Journal of Geophysical Research: Solid Earth*, *122*(1), 387–410. https://doi.org/10.1002/2016JB013132

Yabe, S., & Ide, S. (2018). Variations in precursory slip behavior resulting from frictional heterogeneity. *Progress in Earth and Planetary Science*, *5*(1), 1–11. https://doi.org/10.1186/s40645-018-0201-x

Yamashita, F., Fukuyama, E., Xu, S., Kawakata, H., Mizoguchi, K., & Takizawa, S. (2021). Two end-member earthquake preparations illuminated by foreshock activity on a meter-scale laboratory fault. *Nature Communications*, *12*(1), 1–11. https://doi.org/10.1038/s41467-021-24625-4

Yoon, C. E., Yoshimitsu, N., Ellsworth, W. L., & Beroza, G. C. (2019). Foreshocks and mainshock nucleation of the 1999 Mw 7.1 Hector Mine, California, earthquake. *Journal of Geophysical Research: Solid Earth*, *124*(2), 1569–1582. https://doi.org/10.1029/2018JB016383

Supporting Information for "Physical mechanism for a temporal decrease of the Gutenberg-Richter b-value prior to a large earthquake"

Ryo Ito[1], Yoshihiro Kaneko[1]

[1]Graduate School of Science, Kyoto University, Japan

Contents of this file

Figures S1 to S6

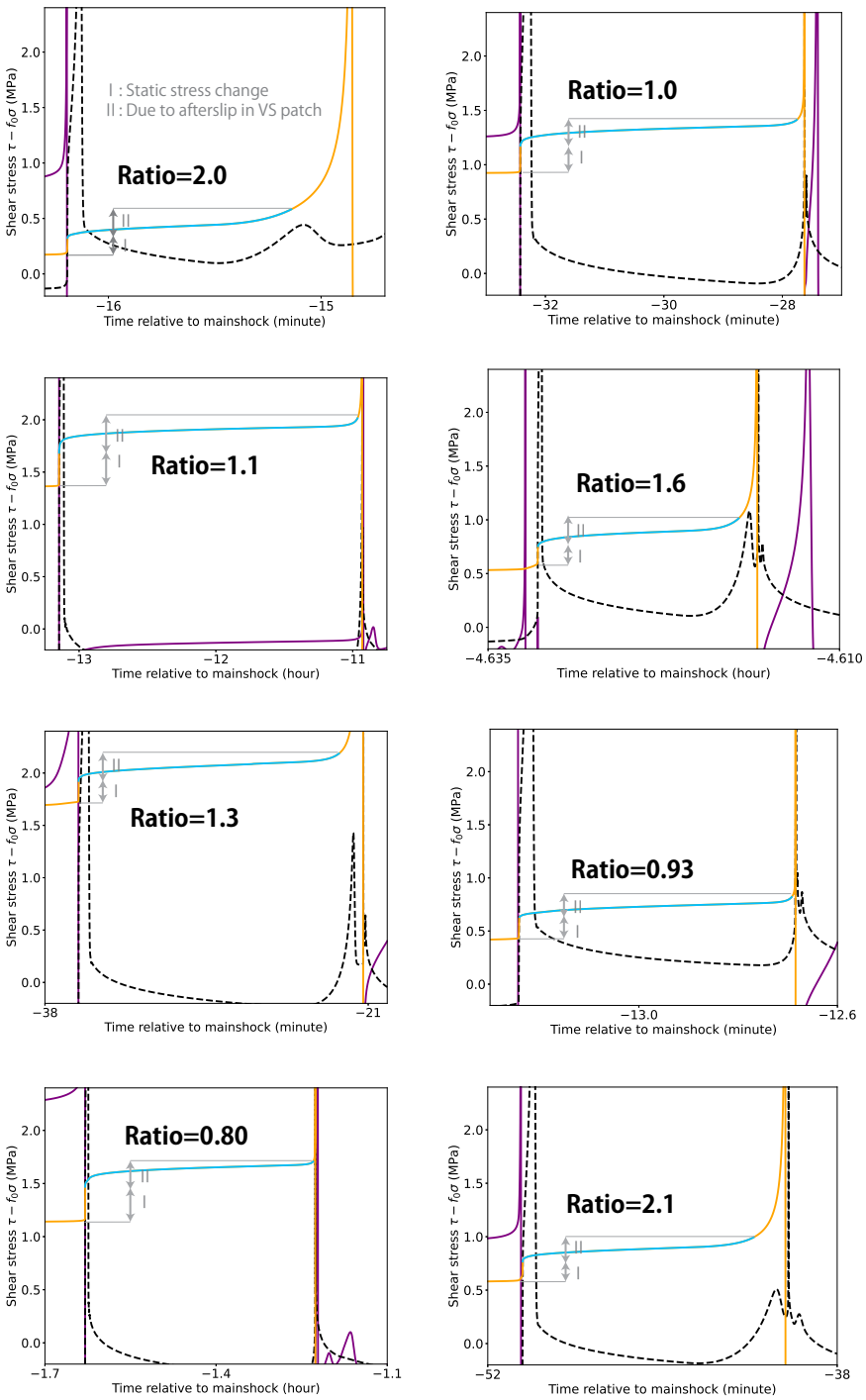

Figure S1. The same as Figure 5 except that randomly chosen, 8 pairs of foreshocks are plotted. The purple line corresponds to shear stress at the VW patch where an earlier foreshock occurred. The black dashed line is shear stress at the center of a neighboring VS patch. The orange and light blue lines indicate shear stress at the VW patch where the subsequent foreshock occurred. For each case, the ratio of the contribution from afterslip in the neighboring VS patch (light blue line) to static stress change is indicated. In triggering another foreshock, the contribution of stress increase within a VW patch due to afterslip in the neighboring VS patches is, on average, larger than that of the static stress change from the previous foreshock.

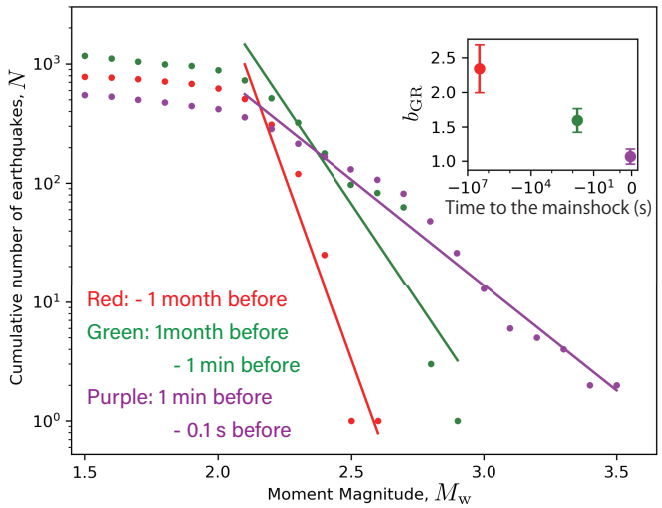

Figure S2. The same as Figure 6 except that non-overlapping time windows are used. Three time windows are used: from the previous mainshock to 1 month before the next mainshock (red), from 1 month before the mainshock to 1 min before the mainshock (green), and from 1 minute before the mainshock to 0.1 seconds before the mainshock (purple). b_{GR} decreases with time regardless of the definition of the time windows.

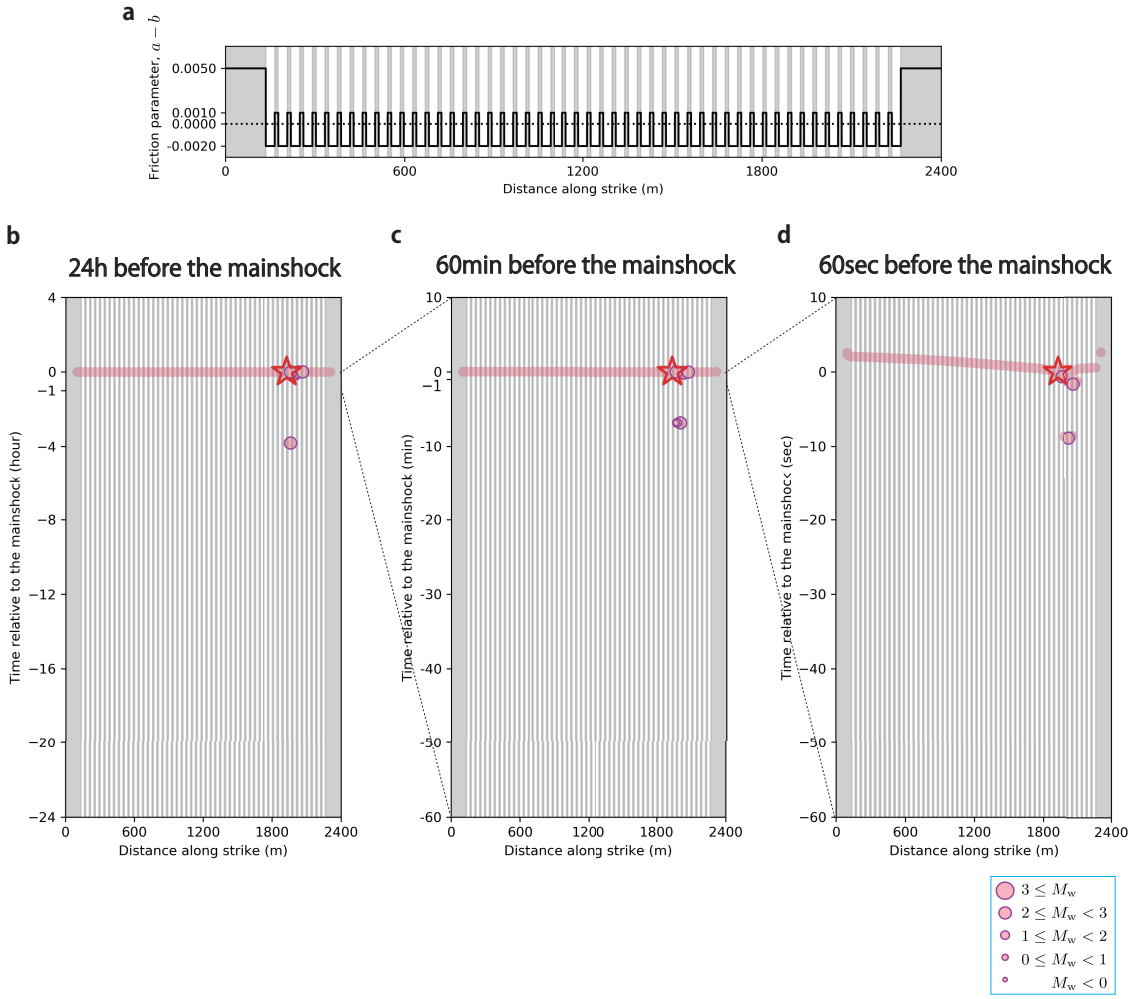

Figure S3. Spatio-temporal evolution of foreshocks in the last 24 hours leading to one of the mainshocks in the case of mainshocks accompanying few foreshocks ($\sigma(a_{vs} - b_{vs}) = 0.050$ MPa). (a) The distribution of rate and state parameters a - b. In this case, $a_{vs} - b_{vs} = 0.0010$, and $a_{vw} - b_{vw} =$ -0.0020. Foreshock distributions over the time periods of (b) 24 hours, (c) 60 minutes, and (d) 60 seconds before the mainshock. Panel (b) is an enlarged image of panel (a), and panel (c) is an enlarged image of panel (b). Purple circles represent epicenters of foreshocks and pink regions represent the rupture area of each event (plotted when slip rate $>$ 1cm/s). Background gray and white areas represent VS and VW patches, respectively. The red star is the mainshock (M_w 4.5). In this case, approximately 10 foreshocks occur within the final 24 hours before the mainshock.

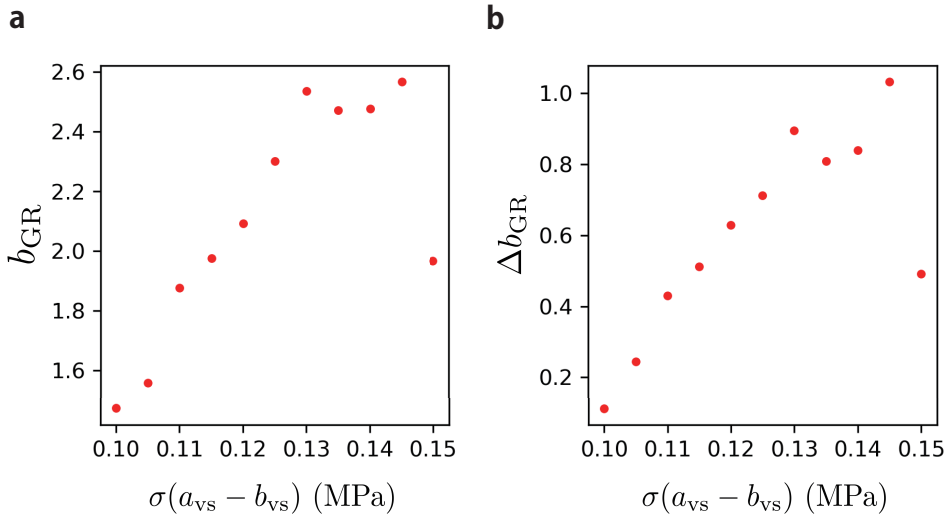

Figure S4. (a) Absolute amount of b_{GR} and (b) change in b_{GR} shown in Figure 10a. Although both b_{GR} and Δb_{GR} vary with $a_{vs} - b_{vs}$, Δb_{GR} has a stronger effect on $\Delta b_{GR}/b_{GR}$ shown in Figure 10a.

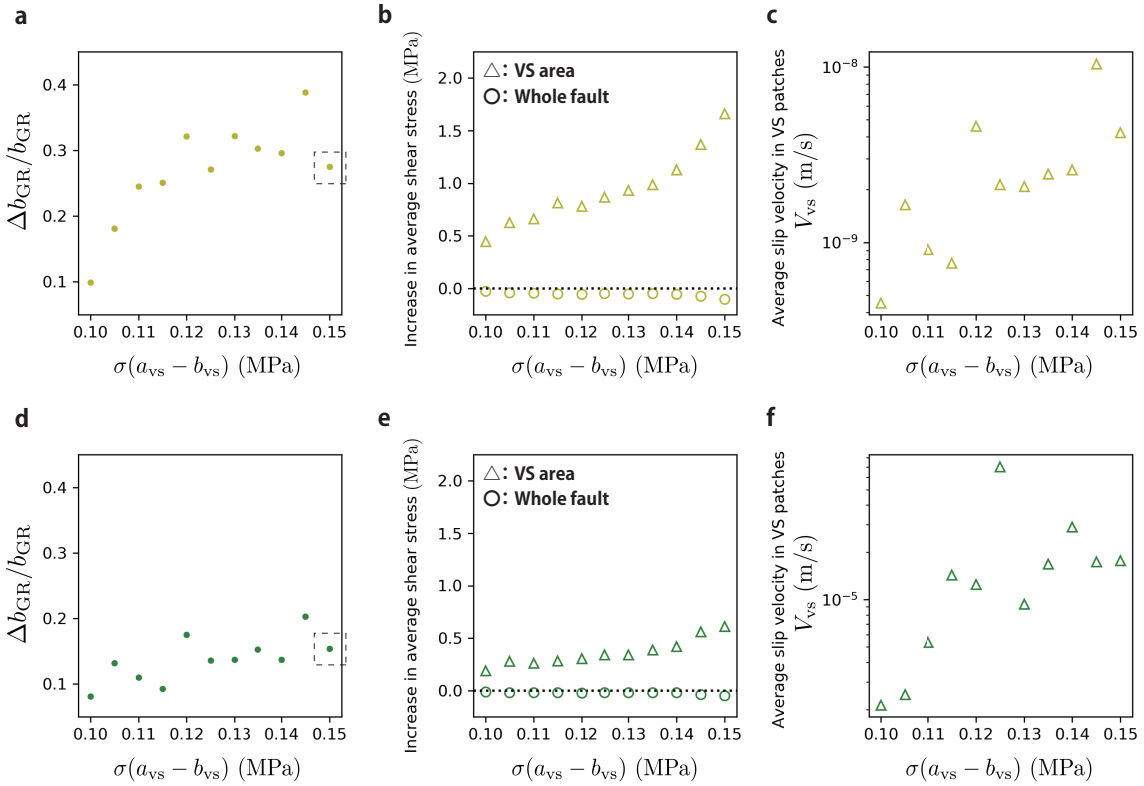

Figure S5. The same as in Figure 10 except that different time windows are used. Each color corresponds to time windows with the same meaning of colors as in Figure 6. The same trends are seen regardless of the chosen time windows.

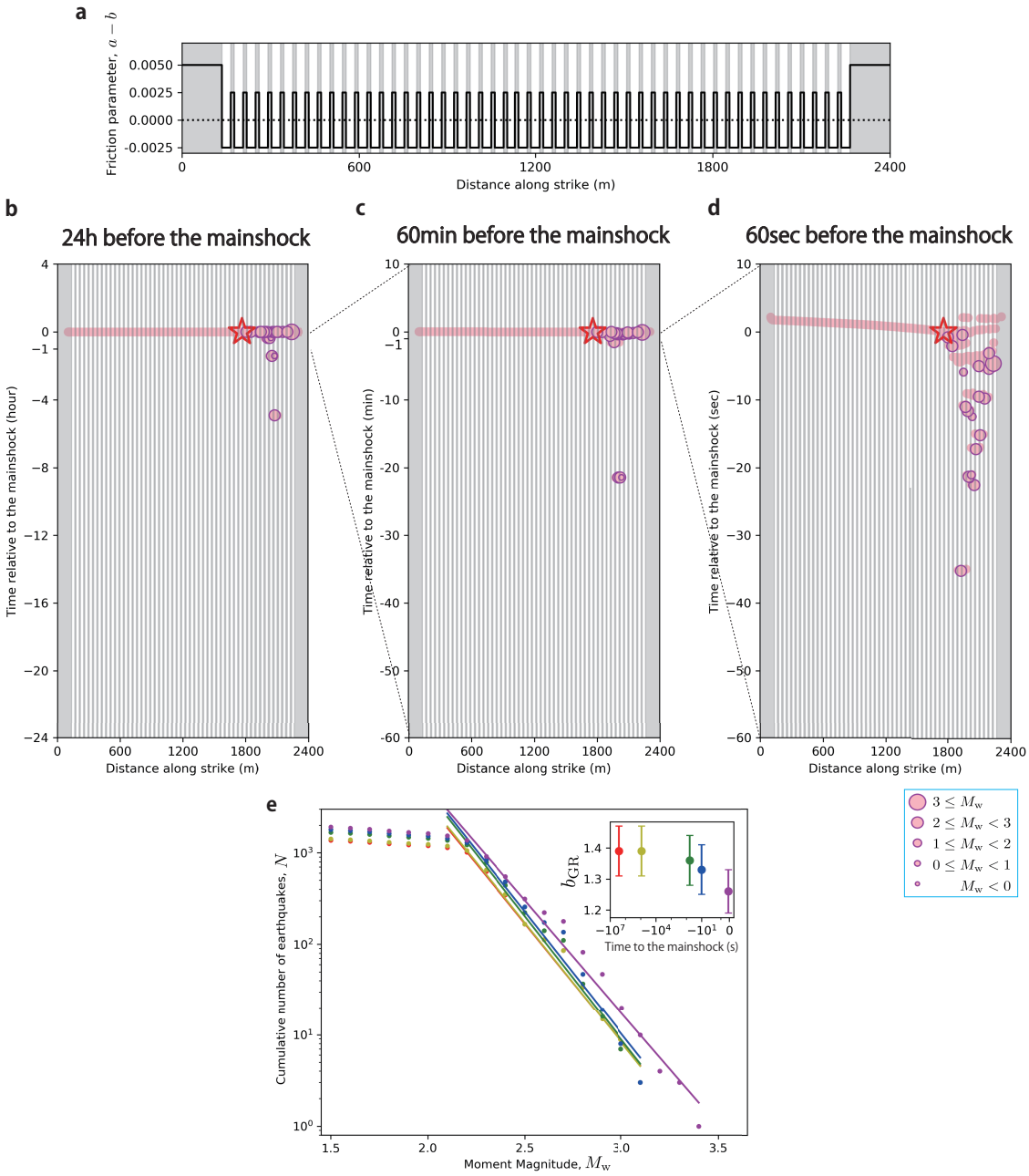

Figure S6. Spatio-temporal evolution of foreshocks and temporal change of b_{GR} in the case with more unstable VW patches. (a) The distribution of rate-and-state parameters a - b. In this case, $\sigma(b_{\mathrm{vw}} - a_{\mathrm{vw}}) = 0.125$, corresponding to 25% increase from the case shown in Figure 4. Foreshock distributions over the time periods of (b) 24 hours, (c) 60 minutes, and (d) 60 seconds before the mainshock. Purple circles represent epicenters of foreshocks and pink regions represent the rupture area of each event (plotted when slip rate $>$ 1cm/s). Background gray and white areas represent VS and VW patches, respectively. (e) Relationship between M_{w} and cumulative number of earthquakes N plotted over five time windows. The inset figure shows the time variation of b_{GR} obtained by the maximum likelihood estimation (Equation 8) with $M_{\mathrm{c}} = 2.1$. The fitting lines are obtained by the least squares method.