Pure Appl. Geophys.

https://doi.org/10.1007/s00024-019-02128-0

Pure and Applied Geophysics

From Natural Stresses in Seismic Zones to Predictions of Megaearthquake Nucleation Zones

Yu. L. Rebetsky[1] and Y.-S. Guo[2]

Abstract—Examines the state in the field of earthquake prediction and modern views of geomechanics on the nucleation of foci of strong earthquakes. From the point of view of tectonophysics, the problem of the mechanism of mega-earthquake occurrence is considered. It is shown that data obtained by tectonophysics methods on the natural stress magnitudes in regions with the strongest twenty-first century earthquakes are the key to its solution. Results of tectonophysical stress inversion define the main area of the mega-earthquake focus as an extended area of low and medium magnitudes of stress. This result corresponds to geomechanical concepts indicating that the state of earthquake focus depends on the Rate and State theory of friction. Numerical geomechanical calculation shows that before the formation of a large-scale brittle fracture, the fault weakens, which manifests in the acceleration of aseismic slip. In real rock, the processes of strength reduction due to increasing aseismic displacement can take years, many decades, and perhaps even centuries. Tectonophysical analysis of natural stress shows that the nucleation of an earthquake can occur both from the boundary of the focus and inside of the focus. In both cases, this small area should be a zone of high stress.

Key words: Earthquake, stress, fault zone, strength, tectonophysics, metastable state.

1. *Introduction*

The epigraph of the presented article could be the question formulated by Kagan (1997a) “If a large earthquake occurs when the stress exceeds the strength of rocks, why do small earthquakes occur over the seismogenic zone all the time”? In the region of a future large earthquake, there are always many *critical state* patches where the rock strength is reached. At the same time, weak earthquakes occurring there do not *trigger* a large earthquake for a long time. In our work we will show why one of these small earthquakes becomes the trigger of a large one.

Earthquake prediction is divided into three types: long-term, medium-term and short-term. Long-term earthquake prediction based on seismological, seismotectonic and geological data does not actually imply the exact date of the earthquake (years–decades before the earthquake) (Fedotov and Solomatin 2017). The medium-term forecast should made a few months before the earthquake. This earthquake prediction is usually based on analysis of the regional seismic regime (Kosobokov et al. 1997). The short-term forecast performed weeks-days before the earthquake can also use seismic analysis. It is believed that the greatest success in earthquake short-term prediction is provided by data on the nature of changes in time for various earthquake precursors (Cicerone et al. 2009). Accumulated experience of medium-term failures of earthquake prediction (Bakun et al. 2005; http://www.mitp.ru/ru/predictions.html) shows that the only task of early and reliable long-term prediction of strong earthquakes is the most achievable. In our work, we will mainly talk about such a long-term earthquake forecast.

In our opinion medium-term prediction failures of strong earthquakes, most of which occur in subduction zones, are associated with the lack of an acceptable understanding of the small earthquake focal zone relationship with the state of the main region from which seismic energy is released. Note, in our work we talk about the forecast of abnormally large—*mega-earthquakes* for Mw > 8.0, the dimensions of which are comparable and even exceed the characteristic dimensions of the crust. We believe that the preparation and implementation of such events in fault zones is different from earthquakes with Mw < 5–6, which better correspond to the ideas

---

[1] Schmidt Institute of Physics of the Earth, Moscow, Russia. E-mail: reb@ifz.ru

[2] State Key Laboratory of Earthquake Dynamics, Institute of Geology, China Earthquake Administration, Beijing, China. E-mail: guoysh@ies.ac.cn

about the formation of brittle cracks obtained in laboratory experiments (Brace and Byerlee 1966; Byerlee 1970, 1978).

The conditions for very strong earthquake occurrences are in the formation of a critical stress state in the region of a future focal zone. It is possible to agree with (Kagan 1997a) that in seismogenic zones the strength *of rocks* is always reached. Probably, in the fault zone there are *special conditions* under which a small earthquake transforms into a very large one.

Brune (1979) suggested that the size of the earthquake focal is determined only in the process of its development, which already implies the *heterogeneity of the rock massif state for* future large earthquakes. Kagan (1997a) also drew attention to the need to explain the reasons not only for the earthquake *nucleation point* in this particular place but also for its completion at a certain site of the fault zone.

What is the scale of these heterogeneities? Are they comparable to the size of the nucleation zone of the mega-earthquake or to the size of its entire focal zone? Hence the need to develop a deterministic approach in the study of the state of seismogenic zones.

In fact, it is difficult to imagine that the huge volume of the earth's crust is relatively evenly and gradually approaching the critical stress state. Such a situation can exist at the micro level, where the size of the inhomogeneities-crystals and grains, corresponds to the size of cracks along their boundaries. It seems that for the earthquake source of hundreds of kilometres, having different scale inhomogeneities of rock elastic and structural properties this is impossible and especially unlikely due to the significant role of fluids in brittle destruction (Rebetskii 2005; Mulargia and Bizzarri 2015), which are often very unevenly distributed in fault zones.

Patches close to a critical state favourable for brittle destruction will always coexist in the region of a future mega-eathquake focal, along with areas in a stable state of sliding or locking (Kanamori and Stewart 1978; Lay and Kanamori 1981). When one of the patches has reached the *static rock strength*, it begins to develop brittle destruction. If the stresses of neighbouring areas are far enough from the brittle strength, then the destruction does not develop further than the first section. Then, only the size of the first section determines the magnitude of the earthquake.

In another case, the emergence of fast, dynamic movement along the fault in the initial phase of brittle fracture causes a change of stress on the developing fault front because the adjacent fault areas have passed into a critical state (Abercrombie and Rice 2005). Then for environment near the initial portion of *dynamic strength* of the rock will be exceeded. Further, the rupture develops in a seismic manner as long as the additional stress arising from its development are sufficient for the brittle destruction of subsequent fracture patches. Under certain conditions, mega-earthquakes can form.

The presence of seismic activity indicates that the fault zone geomedium is near the critical state. In subduction zones, the areas of high seismic activity cover thousands of kilometres. It is important to understand, why mega-earthquakes do not occur constantly under this state. From the analysis it is concluded that it is necessary to distinguish the conditions of formation of the *nucleation zone*—the start of the propagation of the mega-earthquake and the conditions in which the main part of the future focus zone is located. In the first case, the development of fracture occurs in quasi-static conditions corresponding to the limit of *long-term rock strength*. In the second case, the stress increase due to elastic waves from propagation of a seismic rupture leads to the achievement of the limit of *instantaneous rock strength*.

In this paper the existing tectonophysical data on the stress state of the preparation regions for the largest earthquakes of the twenty-first century will be generalized. The principles for deterministic prediction of strong earthquake places will be formulated and the means to search for fault sites where nucleation of large-scale destruction development is most likely will be outlined.

We will show that modern data on natural stress in seismogenic zones, obtained by tectonophysics methods on the basis of seismological data on the mechanisms of earthquake foci for Mw = 4.0–6.5, allow us to identify ways not only to highlight the region of future mega-earthquakes but also the patches of their possible nucleation. When using the stress inversion method, then the range of magnitudes of the studied strong earthquakes determine the averaging scale of stresses as 50–70 km. Therefore, the dimensions of 200 km and more of the earthquake focal region with Mw > 8.0–8.5 allow us to see in detail the nature of the distribution of natural stress that exists before such an earthquake (Rebetsky and Marinin 2006; Rebetsky and Tatevossian 2013). Thus, data on natural stress allows to solve the problem with a long-term forecast, i.e., to allocate seismogenic fault zones where mega-earthquakes are possible. After that, the task of medium-term forecast should be solved. The problem for medium-term forecasts lies in the study and monitoring of processes in the nucleation patches of mega-earthquake foci. Here tectonophysical and geomechanical laboratory experiments can tell when there is a phase of transition from stable creep to dynamic development of displacement on the fault zone (Kocharyan and Novikov 2016; Ma et al. 2012).

Our research is based on deterministic methods for studying the main factor that causes strong earthquakes, i.e., natural stress. Other methods of earthquake prediction are based either on the study of seismic regime features, or earthquake precursors, which by indirect signs can characterize the approach of seismic hazard.

2. *Forecasting Problem*

Research in the field of earthquake prediction developed rapidly in the late 60s, and 70s of the last century. While agreeing in general with the analysis presented in Kagan and Jackson (1991), Nishenko and Sykes (1993), Jackson and Kagan (2011) and Kagan et al. (2012), note that the *Method of Seismic Gaps of First Rank* (Gilbert 1884; Reid 1911) is perhaps the only one capable of providing almost 100% long-term prediction of large earthquakes. In any case, almost all regions of the North-Western flank of the Pacific seismogenic zone, allocated by Fedotov (1965, 1968), experienced strong earthquakes with Mw > 7.7. A significantly different approach for less strong earthquakes is the *Seismic Gap of Second Rank* Method. Here Kagan's statement about the low statistical significance of a positive forecast is absolutely correct.

The Method of Seismic Gaps of First Rank has a significant drawback in that it lacks the ability not only to make short-term but also medium-term forecasts. Until now, we have been waiting for the Avacha earthquake in Kamchatka (Fedotov and Solomatin 2017), the forecast of which was given by Fedotov (1968).

The most famous case of failure to reach a short-term forecast is the Parkfield geophysical polygon for which the place and strength of the future earthquake was known in advance, repeated here many times (Sieh 1978). After its creation early 80s, several medium-term forecasts were made for the new earthquake time. All of them were wrong or so vague that they actually closed the entire period of its recurrence. The event itself, which occurred on September 14th, 2003, was missed (Bakun et al. 2005).

After the forecast failures for the Parkfield geophysical polygon (Kagan 1997b; Langbein et al. 2005; Jackson and Kagan 2006; Bakun et al. 2005) *earthquake prediction* in many leading countries was not priority problem (Kagan 1997a; Kagan et al. 2012; Geller 1997; Geller et al. 1997a, b, 2015; Jordan 2006; Frankel 2013; Kagan and Jackson 2013; Mulargia 2013; Panza et al. 2014). On the other hand, at the end of the last century, new provisions and terms defining geological and seismological features of large-scale brittle fracture development appeared in the physics of the earthquake source. These provisions should influence the problem of earthquake prediction in a certain way.

In particular, it has been shown that in lithospheric plate subduction zones there is a continuous seismic and creep slip that can be interrupted in the so-called *asperity*, where the fracture strength increases (Scholz and Engelder 1976; Kanamori 1978). Data on the degree of asperity are determined based on the complexity of the seismic waves. It is believed that in asperity zones for long periods there is a higher level of stress and increased fracture strength leading to the accumulation of elastic energy (Lay and Kanamori 1981). It was assumed that between the asperities there are patches of aseismic slip. Asperity zones are the most likely places for earthquakes. The position on increased seismic fault zone strength in the future strong earthquake region goes back to the ideas of Reid (1911) and is in good agreement with the theory of stick–slip (Brace and Byerlee 1966).

Lay and Kanamori (1981) shown that different subduction zones correspond to different asperity distributions. Some subduction zones have almost continuous zones of asperity (Chilean), whereas others have little extended asperity (Aleuts, Kurils). To form mega-earthquakes in such subduction zones, the focus must combine several asperities.

Along with the Asperity Model, the Model of Barriers in fault zones was developed (Das and Aki 1977; Mikumo and Miyatake 1978) as a strong earthquake formation source. According to Aki (1984) more of the focal zone is a region of reduced strength where there is a slip. The boundaries of the focal are the patches between two stable and powerful barriers. Inside the focus zone, there are many weak fracture barriers, accompanied by the transformation of elastic deformation energy into seismic energy and heat (friction, plastic deformation).

It is believed that subduction zone earthquakes better correspond to the asperity model (Lay and Kanamori 1980). The barrier model (Corbi et al. 2017) explains well the nature of the aftershock process (Mogi 1962). There are also models, in which barriers between adjacent asperities are assumed to exist.

Theoretical calculations that test these different models are now being developed. They use representations of the dependence of friction on velocity of displacement and state—Rate and State theory (Dieterich 1972, 1979; Ruina 1983). It is shown that the variation of the model defining parameters (sliding velocity and state parameters) can obtain not only the transition from slow slip to a seismic event but also explain other non-standard earthquakes: slow, quiet earthquakes, and so on (Rice 2000; Uenishi and Rice 2003; Mori et al. 2003; Abercrombie and Rice 2005; Rubin and Ampuero 2005; Reches and Lockner 2010). It is also found that both the seismic event and the creep on the rupture are preceded by slip localization in a narrow fault zone.

The localized patch of the original aseismic slip develops and slowly increases in size. Then, in a short time compared to the seismic slip, there is a transition from very slow slip to seismic speeds. Thus, within a certain time before the earthquake, the *asperity* turns into its opposition—the *zone of aseismic sliding with a reduced level of stress.*

The duration of the phase of gradual accelerated sliding can take quite a long time because the maximum level of displacement can reach 20–30% of the seismic. For example, for the Sumatra-Andaman earthquake [maximum seismic displacements in the source of 15–20 m (Ammon et al. 2005)] amplitude aseismic displacement of 3–4 m can occur for 300–400 years, with an average displacement rate 1 cm/year (!) or for 3–4 years at an average displacement rate of 100 cm/year (?).

To explain the failure of earthquake prediction in recent decades, the term *Metastable State of the fault zone* (MSF) has been used (Gol'din 2004, 2005; Sobolev and Lyubushin 2007). It is understood as the special state of the fault zone region before a strong earthquake. The process of a sudden energy-powerful earthquake is similar to the process of a phase transition. Energy accumulation is slow, without visible manifestation and ends with a catastrophic event. The condition for the emergence of a jump-like transition of the medium from one state to another is the presence of the metastable state for part of fault of a future earthquake focus (Rebetskii 2005).

The practice of using the term MSF in earthquake focal physics, shows, that it is understood by the instability state of the earth's crust before fracture, when small external influences (trigger) can lead to a strong earthquake. At the same time, there are practically no physical examples explaining the instability of the geomedium state. Why can the system be in a pre-destructive state for a long time, and what physical processes are responsible for the state change? Often the term MSF for a fault zone is equivalent to the condition of reaching a critical stress state (Sobolev 2011; Kocharyan and Batukhtin 2018).

It will be further shown that there is a tectonophysical justification for MSF for mega-earthquakes. Important to understanding MSF was the introduction of the earthquake nucleation zone concept (Sobolev 2011; Kocharyan and Batukhtin 2018). In fact, the nucleation zone may be understood as a small earthquake that is the trigger for a strong earthquake.

Examples of strong earthquake forecast failures, show that the seismic regime of future earthquakes can indicate that it is prepared for a strong earthquake but that may not occur for many years [the global test of earthquake prediction http://www.mitp.ru/ru/predictions.html on the base of algorithm (Keilis-Borok and Kossobokov 1990; Kossobokov et al. 1997; Davis et al. 2012)]. As noted above, this may be due to the strong heterogeneity of the main region of the future mega-earthquake focal, in which there are too many zones far from the critical state. However, it can also be related to the patch of earthquake nucleation. If power trigger earthquakes are not enough to activate the dynamic prepared zones of the main focal, then a strong earthquake does not occur. Since statistical methods study the main focal and ignore the small nucleation zone, most part of their results of medium-term prediction may be false alarms.

In our view, the problems of medium-term forecasting lie in the study of the state of nucleation zones. It is in the studies of these zones that the results of tectonophysical and geomechanical modelling of brittle fracture should be used to greater extent. Ma et al. (2012), Ma and Guo (2014) provided a new model to investigate the evolution stage before fast fault instability, fault meta-instability (sub-instability). In stress–strain curve, the stage begins at the peak stress point, ends at the fault dynamic instability.

Further, this paper will present the results from the study of natural stress states in formation regions of the strongest twenty-first century earthquakes that made it possible to formulate tectonophysical criteria to identify regions of MSF and forecast of nucleation zones—fault patches that trigger earthquakes.

3. *Tectonophysics of Fault Zone Stress States*

3.1. *Coulomb Stress*

In Russia, M. V. Gzovsky since the end of the forties of the last century developed a method for reconstruction of natural stress (stress inversion), based on the allocation of conjugate pairs of shear cracks (Gzovsky 1954a, b). These approaches were similar to the approaches of Anderson (1951) and made it possible to determine only the orientation of the axes of the principal stresses. Later, along with data of geological natural deformation indicators, seismic indicators of deformation mechanisms for earthquake foci were used for stress analysis. In these works, tectonic stress was taken to be the orientation of the *P* and *T* axes of individual earthquake focal mechanisms. It is not true. It is known (see (Kenzie and Dan 1969; Kostrov 1975) that the *P* and *T* determine the orientation of the axes of the unloading principal stresses taken in the area surrounding the seismic discontinuous displacement.

Gzovsky (1957, 1975) set the task of creating tectonophysical criteria for seismicity. It was necessary to use data on natural stress and surface movements to learn how to rank fault zones according to their degree of danger. It was originally thought that the areas of high magnitude of maximum shear stress are dangerous:

$$\tau = 0.5(\sigma_1 - \sigma_3), \tag{1}$$

where $\sigma_1 \geq \sigma_2 \geq \sigma_3$ are the principal stresses. In our work, the stress sign rule adopted in classical mechanics will be used, i.e., tension is positive. In this case, σ_1 and σ_3 are, respectively, the minimum and maximum compression. Because the deviatoric components of these stresses:

$$s_i = \sigma_i - \sigma, \quad \sigma = -p = \sigma_i/3, \tag{2}$$

are, respectively, the greatest tension and compression, in the future we will use the terms principal tension and principal compression. In expression (2), σ is the mean stress, and p is the isotropic pressure. Since tectonophysics in the 50s and 60s of the last century was not able to calculate the magnitude of maximum shear stress, it was proposed to use data on horizontal gradients of vertical movements for its evaluation. However, this approach, along with the successful results of the allocation of hazardous crustal regions also failed.

Currently, after summarizing the results of experiments on the destruction of rock samples (Brace and Byerlee 1966; Byerlee 1978) it is known that shear fracture formation and rock brittle fracture correspond to Coulomb stresses (Karato 2008; Rebetsky and Polets 2018):

$$\tau_C = |\tau_n| + k_f\sigma^*_{nn},\ \sigma^*_{nn} = \sigma_{nn} + p_{fl}, \quad \sigma^*_{nn} < 0, \quad (3)$$

which depend both on the deviatoric (1) and isotropic component (2) of the stress tensor. In expression (3) τ_n and σ_{nn} are the shear and normal stress on the crack plane, σ^*_{nn} are effective stress, k_f is the coefficient of internal (for whole samples) and surface friction (for samples with existing cracks), and p_{fl} is the fluid pressure in rock crack—pore space, which reduces the compression on the crack from compression in the solid matrix.

A certain level of deviatoric stress is necessary for brittle fracture formation, but if a high level of normal compression stress acts on the fracture plane, then it may not occur. In Fig. 1 an example of a Mohr diagram with two conjugate planes (equivalent to the nodal planes of focal mechanisms) is shown, one of which is a seismogenic discontinuity—fault. For ease of understanding, two orthogonal planes are taken, the stress states of which lie on a large Mohr circle. For them, the magnitude of shear stress is the same $\tau^1_n = \tau^2_n$, but the normal stresses are different $\sigma^1_{nn} > \sigma^2_{nn}$ (less compression—σ^1_{nn}), i.e., for the first crack the compression is smaller than for the second. Therefore, $\tau^1_C > \tau^2_C$ and more likely to form a fault for the first nodal plane. If we assume that the line of minimum friction resistance passes above point τ^2_n, σ^2_{nn}, the second nodal plane cannot be realized in the form of a seismogenic fault. Moreover, the second example for conjugate planes (points 3 and 4 in the Mohr diagram) shows that the Coulomb stresses are greater for a plane with lower level of shear stress ($\tau^1_C > \tau^3_C = \tau^4_C$).

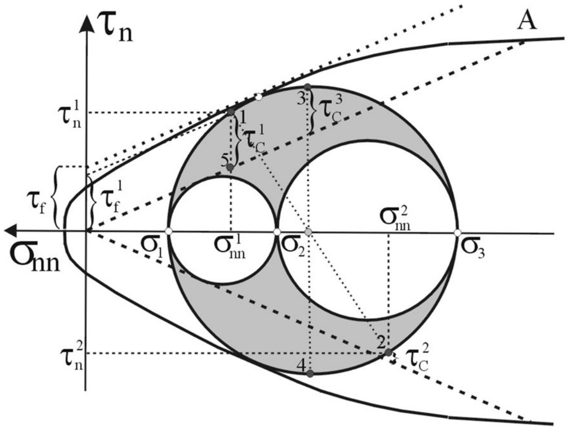

Figure 1
Analysis of stresses on a Mohr diagram. The envelope line is the limit of brittle strength of the crust, dash-dotted lines minimum friction resistance on the rupture. Points 1, 2 and 3, 4 correspond to the stress states of two pairs of orthogonal planes. Positive values of the effective normal stresses are plotted (σ^*_{nn}) to the left (explanation in the text)

The area between the curve of rock brittle strength and the line of minimal resistance to dry friction, is a *zone of brittle fracture* for faulting rock. If the cohesion strength, $\tau^i_f(\tau^i_f < \tau_f)$ of the existing crack becomes equal to the Coulomb stress, τ^i_C corresponding to it, this crack will be activated again. We stopped at this simple example in such detail because at present in the seismology and physics of earthquake foci there is still an idea of increased danger for regions with a high level of deviatoric or maximum shear stress. At the same time, laboratory experiments and the results of tectonophysical analyses of natural stresses suggest the opposite. As an example, here is a phrase "…the fractal pattern of the earthquake fault geometry is due to the self-organization of the fault under high lithostatic and tectonic shear stresses." Earthquake prediction methods based on the assessment of the stress accumulation time period also proceed from the hypothesis of the deviatoric stress level as the cause of brittle failure or earthquake. Examples are stress-accumulation models such as the time-predictable and slip-predictable schemes proposed by Shimazaki and Nakata (1980).

From the energy point of view, medium and even low-level stresses are more dangerous, as was noted in the monograph by Rice (1982): "strong earthquakes should not fall into the area of high stress levels, because here on the planes of the gap there is a high level of friction, to overcome which will take most of the released energy."

The Mohr diagram allows one to graphically estimate the level of stress relieved since for shear cracks, the normal stresses before and after crack activation ($\sigma^1_{nn} = \sigma^5_{nn}$) remain unchanged (Osokina 1987). For Fig. 1, point 5 is the new state of the point 1 after crack activation is simplified to coincide with the minimum static friction resistance. In this case $\Delta\tau^1 = \tau^1_f$. In reality, it is slightly larger because the dynamic friction on the crack during its motion is lower than the static (Mori et al. 2003).

Note that the relieved stresses for the activated cracks are the highest for the average stress level (Fig. 1). This characterizes this range of stress state as the most effective for brittle fracture. Where dry friction moves closer to the curve of brittle failure tensile strength it is least efficient (point A Fig. 1), and after their intersection it is impossible.

3.2. *The Main Regularities of Natural Stresses of Seismic Focal Zones*

Currently, tectonophysics has advanced significantly in the study of natural stress. In the early 90s of the last century the *Method of Cataclastic Analysis of discontinuous displacements* (MCA) (Rebetsky 1996; Rebetskii 1997) had an algorithm that could only determine the shape and directions of the stress ellipsoid principal axes. At that time the MCA algorithm was similar as well to methods of O. I. Gushchenko, J. Angelier, and J. Gephard. In this part the MCA is similar to the methods Gushchenko (1975, 1996), Angelier (1979, 1990). However, already at the beginning of the new century (Rebetskii 2003) an in frame MCA developed an algorithm ratio for calculation of stress tensor spherical and deviatoric components, and the determination of the stress magnitudes, array brittle strength (cohesion) and fluid pressure (Rebetsky 2007, 2009).

It is important to note that the basic data on the principal stress axes are obtained on the basis of seismic deformation indicators. The very possibility of obtaining the deviatoric and spherical components of the stress tensor in the MCA is associated with the reduced Mohr diagram in the analysis of stresses on seismic focal planes, as well as additional data on the magnitudes of stresses dropped in the foci of the strongest earthquakes in the region under study. In this part, the MCA algorithm is similar to the ideas expressed in the works of Angelier (1989) and Reches (1983), which developed methods for the evaluation of stress magnitude according to the combination of shear cracks.

The main terms and algorithms of the MCA (Rebetsky 2007; Rebetsky et al. 2017; Rebetsky and Polets 2018) are: (1) the creation of homogenous sets of earthquakes with data on focal mechanisms that satisfy the principle for the dissipation of internal mechanical energy; (2) the approximation of brittle fracture zone by strip on the reduced Mohr diagram. In this strip, stress analysis is performed for mechanism foci from a homogeneous set of earthquakes; (3) selection of the realized earthquake nodal plane based on the hypothesis of maximum Coulomb stress (Fig. 1); (4) a presence in a homogeneous set of earthquakes with a minimum strength corresponding to the friction strength in the fracture strip; and (5) evaluation of the static stress drop for each earthquake from a homogeneous set based on the positions of elastic cracks (normal stress on the crack plane is the same before and after its activation/occurrence).

The main difference between the MCA and the statistical analysis methods of the P and T axes of the set of earthquake foci mechanisms (Zoback 1992; Heidbach et al. 2010) is the physical limitations (point one of the previous paragraph) imposed on earthquakes from a homogeneous set, on the basis of which the axes of the principal stresses are determined, which allows further assessment of the normalized stresses on the Mohr diagram (paragraph two of the previous paragraph) since all of the mechanisms of earthquake foci should fall into the brittle fracture band. In the methods of M. L. Zoback this is not possible because the set on which the axes of the principal stresses are determined, includes all earthquakes that fall into the predetermined stress averaging window in advance.

In tectonophysical studies of seismic regions carried out at the beginning of the 0 years using the results of reconstruction of natural stress (Rebetsky and Marinin 2006; Rebetsky et al. 2012), it was shown that seismic fault zones correspond to a sufficiently mosaic structure of stress magnitude distribution, characterizing the multi-scale heterogeneity of the stress field. Such mosaic distribution of stress is due to the specifics of the structural-physical states of different parts of the faults (Rebetsky 2006) and also with the accuracy of stress estimates.

Figure 2 shows locations of the inland Altai–Sayan orogeny area and two regions of active continental margins for East Asia and Indonesia with examples of stress inversion results by MCA (Rebetsky et al. 2012, 2016a, b; Rebetsky 2009; Rebetsky and Tatevossian 2013; Rebetsky and Marinin 2006). The results themselves are in the form of the spatial

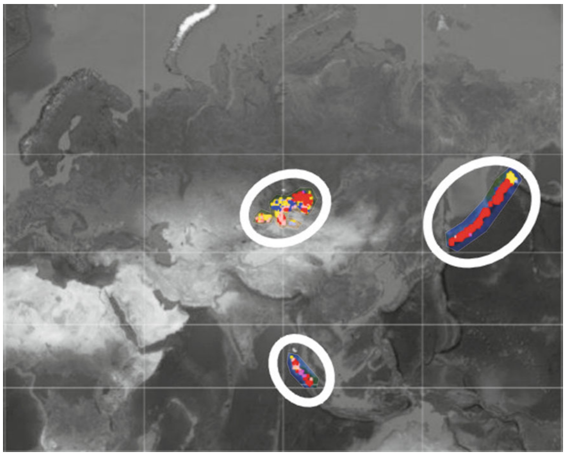

Figure 2
Locations of regions with examples of stress inversion results by MCA (see Figs. 3, 4 and 5)

distribution of various stress tensor components: the stress axes of the greatest horizontal compression (S_H^*), the underthrust shear stresses τ_z on the horizontal planes, and the normalized maximum shear stresses (τ/τ_f) are presented in Figs. 3, 4 and 5. As you can see, there are zones of high and low maximum shear stress for the regional averaging scale (50–70 km). These stresses largely determine the resistance to slip on faults. Zones of increased stress can be interpreted as regions of increased strength, i.e., as an asperity. Zones of decreased stress can be interpreted as regions of weakening and the expected increased sliding velocity at faults. It is important to note that the identified regions of high stress have characteristic dimensions from the first tens to the first hundreds of kilometres.

Estimates of the stress magnitudes confirmed the statement made in Lay and Kanamori (1981) about the different nature of the asperity distribution in subduction regions. For the South American and Japanese zones, the length of the asperity sections was 600–800 km, for the Western Flank of the Sunda Arc it was 300–400 km, and for the Kuril-Kamchatka, 50–100 km. Note that in all cases, except for the Japanese seismic region, we are talking about stress averaged for the entire power of the conditional crust (30–40 km).

Already the first results of the stress magnitude estimates allowed one to pay attention to the fact that the critical stress states located in different parts of the Mohr diagram have different mechanisms of mechanical energy dissipation: (1) strong earthquake—large scale brittle destruction of the earth's crust; (2) a large number of medium and weak earthquakes—cataclastic (pseudoplastic—mech.) or a quasi-brittle flow along an extended zone of the crustal fault; and (3) quasi-static creep along the fault—viscosity or quasiplasticity flow (due to the microcracks on the level of grains) (Rebetskii 2005; Rebetsky 2006, 2007; Rebetsky et al. 2012, 2016a, b; Rebetsky and Tatevossian 2013).

Our previous research has established:

1. Stress magnitudes are extremely inhomogeneous in the earth's crust, their difference in one seismic region can reach 1–1.5 orders of magnitude. Zones of increased stress levels are consistent with the definition of asperity (scheme of the maximum shear stress Figs. 3, 4, 5).
2. There is a certain range of ratios between the effective pressure and the maximum shear stress of 0.5–2 associated with the requirement to meet the Coulomb–Mohr criterion for brittle fracture $k_f = 0.5–0.7$).
3. The estimate of real fractured massif cohesion of 0.1–5 MPa corresponding to the averaging scale of calculated stress in the first tens of kilometres, is much lower than the strength of whole samples of crystalline rocks of 5–10 cm in size (10–50 MPa).
4. The brittle strength of the intracontinental orogeny arrays is 3–5 times higher than the strength in the lithospheric plate subduction zones, which determines the higher level of deviatoric stress and effective pressure in the continental crust;
5. The level of maximum shear stress in the subduction zones generating the strongest earthquakes ranges from 0.3–0.5 to 5–10 MPa.

During the stress inversions in different seismic regions, important regularities were also obtained that carry information about the mechanisms of their deformation deformation (Rebetsky and Tatevossian 2013; Rebetsky 2015; Rebetsky et al. 2012, 2016a, b). In the development of the seismic hazard problem, it was found that relatively strong regional earthquakes (Mw = 6.0–7.0) rarely fall into

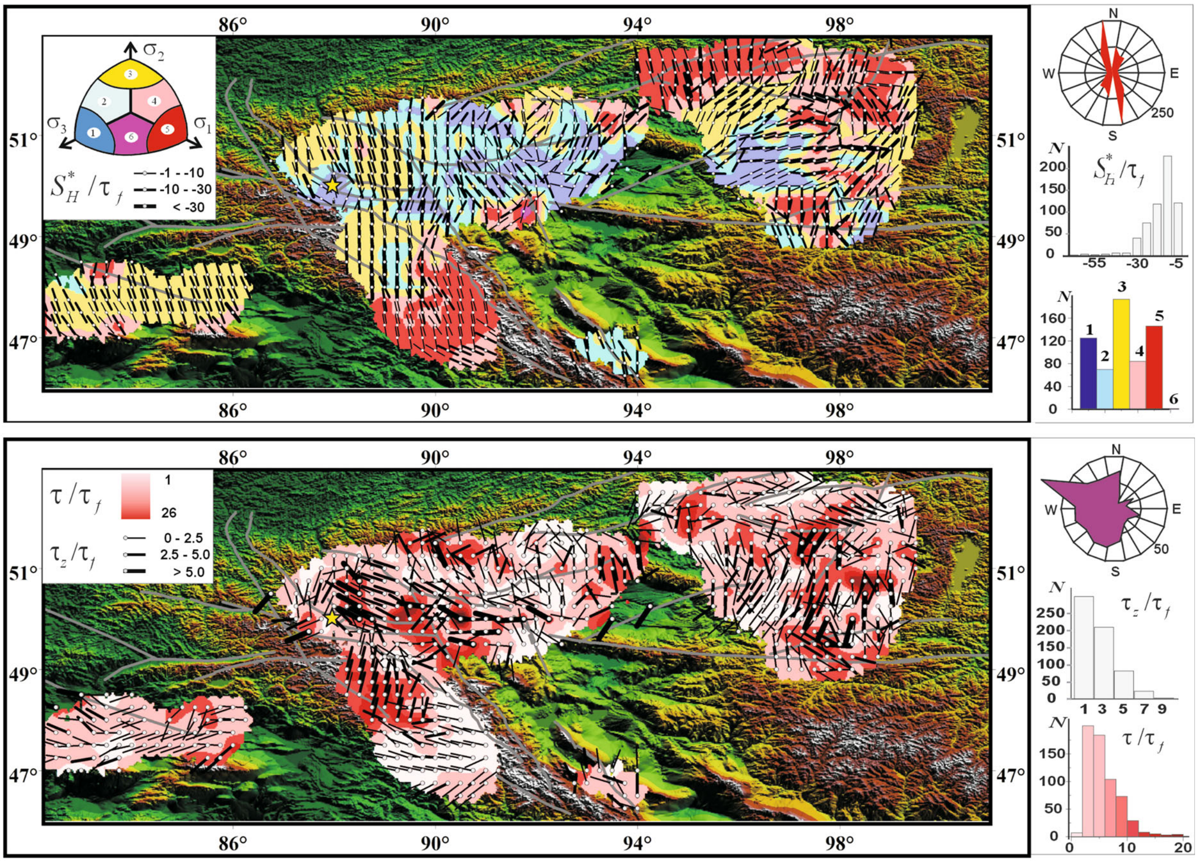

Figure 3

Parameters of stress state for the crust of the Altae–Sayan: **a** axes of the greatest horizontal compression, S_H and geodynamic type of stress state; **b** axes of the underthrust shear stresses, τ_z and normalized maximum shear stress, τ/τ_f. Diagram in the upper left corner **a** shows geodynamic shape of stress state by zenith vector: 1, horizontal compression; 2, horizontal compression and shear; 3, horizontal shear; 4, horizontal tension and shear; 5, horizontal tension; 6, vertical shear. Rose and rectangle diagrams show the predominant trends of strike of these stress axes and predominant range of geodynamic regime of stress state and normalized stresses

the zone of high level of maximum shear stress (Figs. 3, 4, 5).

Their epicentres are usually located in areas of average level of stress or in areas of its high gradient. This empirically observed fact coincides with Rice (1982).

Fundamentally important for creation of tectonophysics model mega-earthquake was the reconstruction of natural stress in the source of the great Sumatra-Andaman earthquake (SAE) 2004, Mw = 9.1 (Fig. 4), obtained according to the Global CMT catalogue focal mechanisms of earthquakes with Mw > 4.3 occurred in the period from 1978 to December 2004. It was found that along the seismic region of the SAE there was a significantly inhomogeneous stress state. To the south of the SAE propagation start, there was an increased level of stresses, and to the North, it fell sharply. The absence of stress data on part of the seismic region is associated with a small number of earthquakes with data on the mechanisms of earthquake foci, which indirectly also indicates a low level of Coulomb stress. Thus, the development of the SAE source occurred not in the high but failed low stress region, where there is a reduced compressive pressure on the

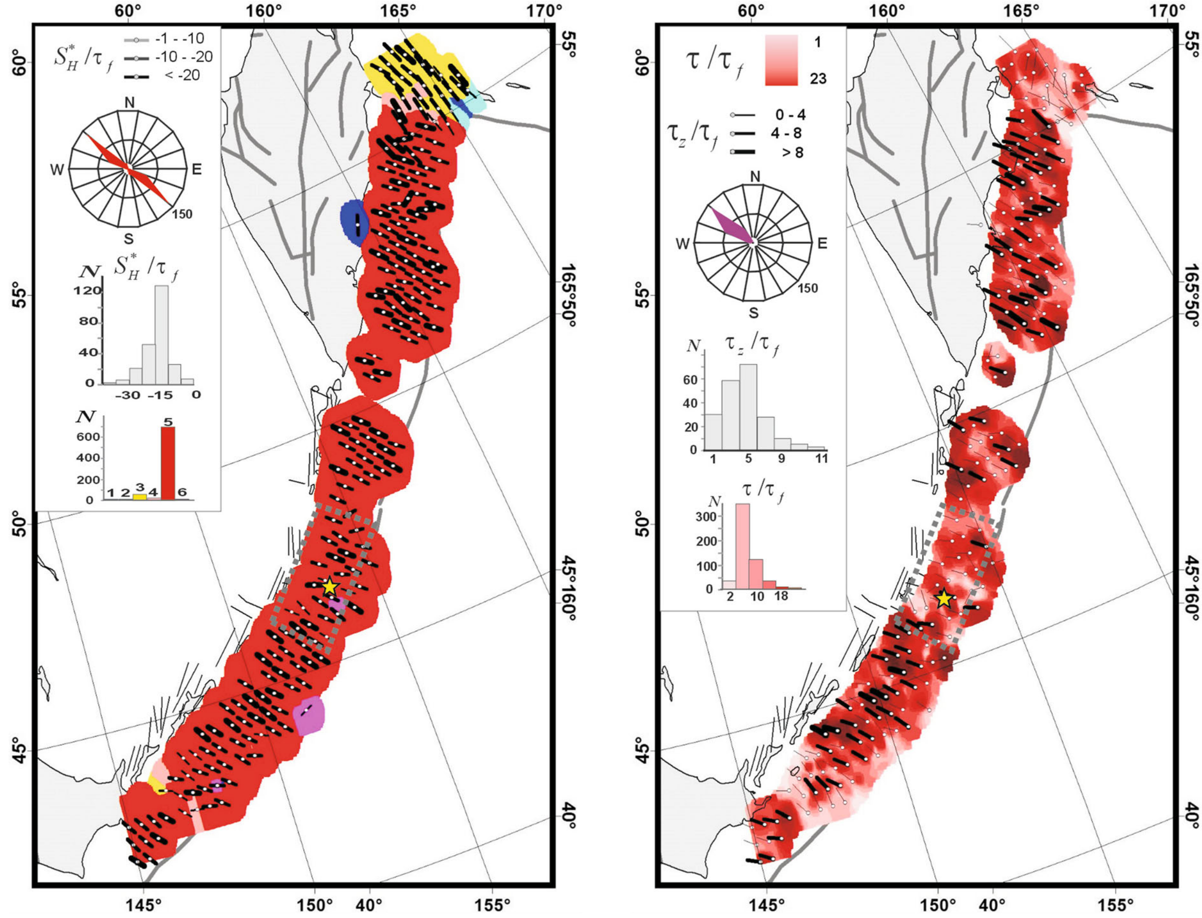

Figure 4

Parameters of stress state for the crust of the Western Flank of the Sunda Arc: **a** axes of the greatest horizontal compression, S_H and geodynamic type of stress state; **b** axes of the underthrust shear stresses, τ_z and normalized maximum shear stresses, τ/τ_f (see Fig. 3)

fault and therefore brittle destruction allows it to release more energy (see Fig. 1).

Data on stress for the Kuril subduction zone before the Simusher earthquake 2006, Mw = 8.3 (Fig. 5) also showed that the focus was located in the region with mean level of stresses and was limited from the North-East and South-West by areas of crust with high stress level. For the Chilean Maule earthquake of 2009, Mw = 8.8, the focus was also located in the zone of low and medium stress, and its northern boundary was the southern limit of the extended zone of high effective pressure (Rebetsky and Tatevossian 2013), i.e. confirming the results of stress inversion for failed Sumatra-Andaman earthquake.

3.3. *Tectonophysical Model of the Metastable State of Faults*

New tectonophysical data on the stress magnitudes obtained in the most recent years in the areas of preparation for anomalously strong twenty-first century earthquakes (Rebetsky et al. 2012; Rebetsky 2015), allow us to move to the creation of a geomechanical model of mega-earthquake foci and filling the physical content of the term MSF.

Since all the data on the state of earthquake foci, obtained by us previously are associated with seismogenic regions of active continental margins further, the term fault will be to understand the seismic zone of tectonic plate boundaries.

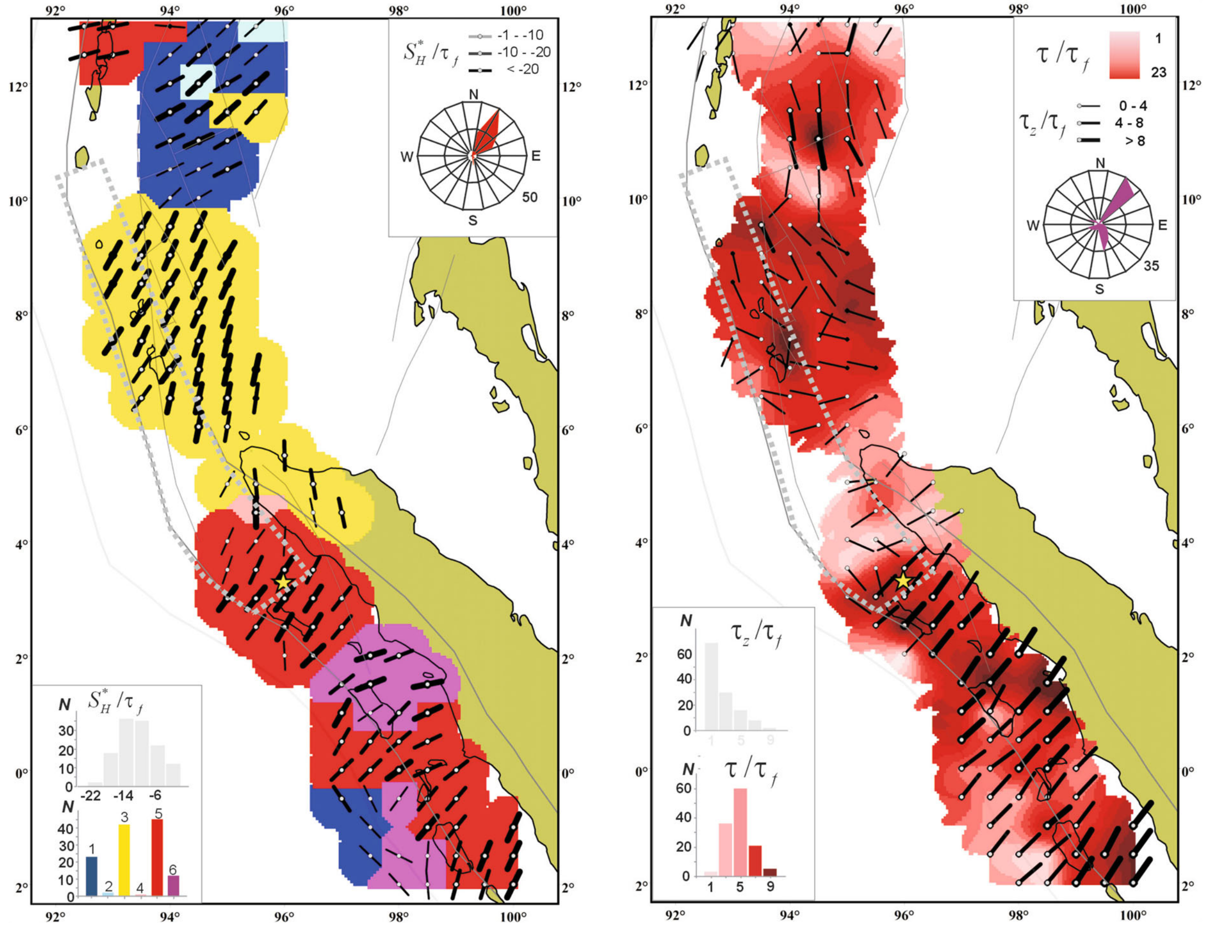

Figure 5

Parameters of stress state for the crust of the Kuril and Kamchatka: **a** axes of the greatest horizontal compression, S_H and geodynamic type of stress state; **b** axes of the underthrust shear stresses, τ_z and normalized maximum shear stresses, τ/τ_f (see Fig. 3)

The formulation of the MSF tectonophysical model is related to the heterogeneity of deformation (stress) of a constantly manifesting seismogenic fault within the future earthquake focal and its immediate environment. The results of tectonophysical stress inversion show that the foci of a number of mega-earthquakes with a magnitude greater than Mw > 8.5 (Rebetsky and Marinin 2006; Rebetsky 2007, 2009; Rebetsky and Tatevossian 2013) are allocated in the field of effective pressure and stress as an region of middle or low from their level. Within this region of the fault the stress is distributed quasi-uniformly, and outside the stress increases sharply. Thus, the limitations of the future foci of mega-earthquakes are always the asperities. Patches of increasing stress inside mega-earthquake foci may be interpreted as weak barriers (Aki 1984). A large area of seismogenic faults with *low or middle stress level* (near 300 km) will be called *the first condition for preparation of mega-earthquakes* (zone Lq on the Fig. 6). Note that here we are talking about an averaging scale of stresses corresponding to lateral 30–70 km and the entire crust 40–60 km in depth.

Note that the regions of low stress level can be considered as former asperities, which in the course of their development experienced peak stress and moved to the final stage of increased slip. Recall that according to modern concepts of the Rate and State

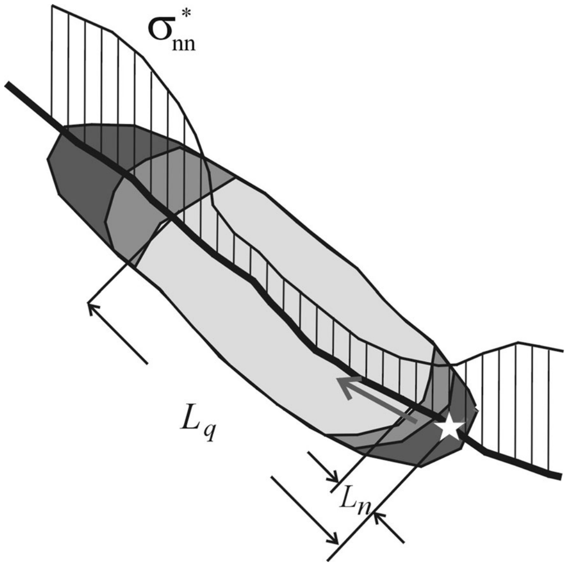

Figure 6
The model of the mega-earthquake focal on the works of Yu. L. Rebetsky. The zone of lower effective pressure (σ^*_{nn}), L_q is focal length of possible future earthquake. An asterisk shows the beginning of the propagation earthquake. L_n is the zone of the maximum stress gradient (trigger earthquake)

theory of strength (Dieterich 1979, 1992), as soon as such a zone exceeds the limit magnitudes of displacement, strength drops sharply.

On the other hand, in a number of studies it was suggested that the low stress level regions are long-existing areas of the fault zone that arose naturally in the process of evolution.

For the Sumatra–Andaman mega-earthquakes, it is known that the seismic rupture propagated from the southern boundary of the focus to the north. These earthquakes began at the northern edge of the zone of high stress and effective pressure and propagated to the region of low magnitude. Thus, we can say that the zone of nucleation for the Sumatra–Andaman mega-earthquake was located on the patch where the regional scale stress had a high gradient. Seismological data (Lay et al. 2005) show that in this patch the rupture propagated at the highest rate. For the Maule and Simushir mega-earthquakes, the *nucleation zones* (hypocentre) were located inside the foci and there were sites of local regional stress increase in the immediate vicinity. The nucleation zone—patches of high stress gradient (Ln on the Fig. 6) is the *second condition for preparation of mega-earthquakes*. This patch is the area where the foreshock activity should occur.

Patches of fault zones with a high level of stress are where fractures begins occur, but it will certainly spread to low stress zones. This conclusion can be drawn from analysis of the Mohr diagram (Fig. 7). In areas with high stress levels, the magnitude of the stress drop is small (the vertical segment of grey colour in the Mohr diagrams Fig. 7). Consequently, such destruction areas do not lead to the release of large mechanical energy. On the contrary, if the stress level is low or medium, the stress drop is large. Rupture propagation from high-stress to medium-and even low-stress regions leads to an increase in the magnitude of the stress drop and, consequently, to an increase in the released mechanical energy and a more efficient replenishment of the rupture side motion kinetic energy (Fig. 7).

The stress analysis shows that the nucleation zone of mega-eathquake is small and equivalent to the size of the earthquake focal with Mw = 7–7.5 (50–100 km). The nature of the stress gradient in the transition from nucleation plays a large role in the fact that any earthquake in this zone will trigger an mega-earthquake to MSF. According to (Aki 1984) the presence of a sufficiently homogeneous stress gradient in the nucleation zone makes it possible to increase the rupture side motion kinetic energy in the event of an earthquake due to the absence of paths of increased strength that prevent the destruction of weak barriers (Aki 1984).

To start a mega-earthquake, there must be no local areas of high stress in the region of high stress gradient that should be considered weak barriers. Having them can prevent a triggering earthquake from turning into a mega-earthquake. It should be understood that the results of tectonophysical stress inversion give us data on the regional scale average stress (Fig. 7). This scale of averaging is sufficient to allocate a section of low stress with a length of 300 and more kilometres. However, it’s bigger than the typical size of the 30–70 km earthquake nucleation zone. Therefore, stresses of regional scale does not allow us to characterize the zone of the stress gradient. To understand what is happening in this trigger zone, it is necessary to have data on *stresses of local averaging scale* (10–15 km). For Fig. 7 two

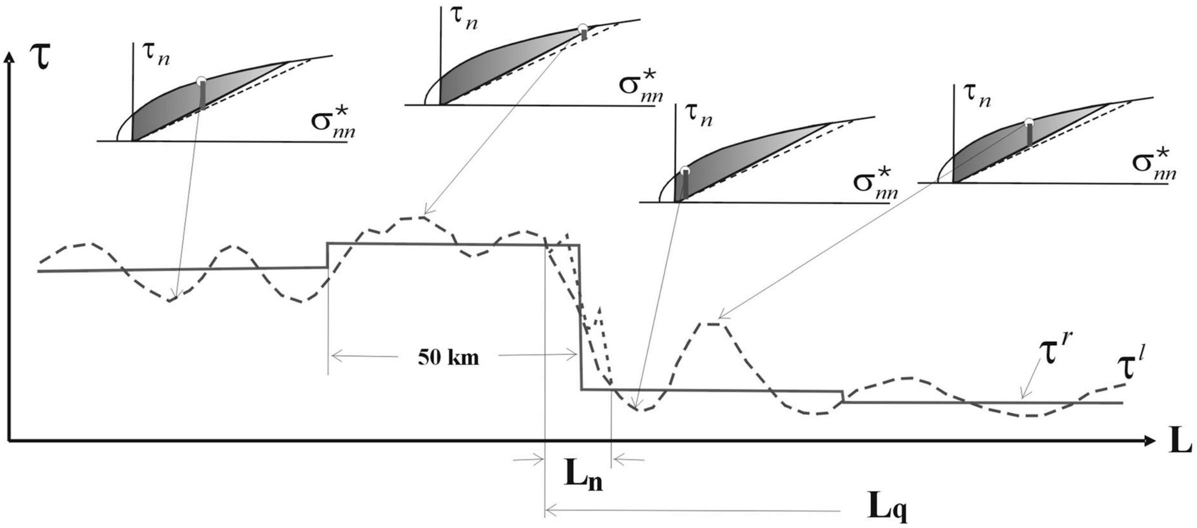

Figure 7
Inhomogeneities of stresses of the first—the second ranks in the field of preparation of the mega-earthquake. The continuous piecewise smooth line is the stress of the regional level, the first stress rank obtained from tectonophysical reconstruction with averaging of 50 km for four points, two of which (on the right) lie in the area of the future focal of a strong earthquake. Dash-dotted line-local of the second rank of stress, determine the level of the stress gradient in the zone of nucleation of earthquake. Dotted line-sub-local level of the third rank of stresses, determine the maturation of the stress gradient zone. The presence of local peaks of high effective pressure is considered as an obstacle to the emergence of a foreshock, initiating a strong earthquake. Above shows the Mohr's diagrams with the stress drops (vertical sigment) for the stress state of the local level

variants of stress distribution in the transition patch from high to low stress regions are shown. In one, the stress gradient is uniform (dash-dotted line), and in the second, it is not (dotted line). For last variant there are weak barriers that hinder the effective development of the trigger rupture. Thus, the homogeneity of the stress gradient in the nucleation zone shows its readiness for a trigger earthquake.

The next position of the tectonophysical model of MSF is also associated with the *heterogeneity of stress on a local scale*, but in the throughout focus of the mega-earthquake. It is likely that within the fault zone in the future focal of mega-earthquakes, there may be relatively small areas that are not ready for brittle destruction. These instabilities give rise to significantly larger stress heterogeneity than in the zone of nucleation. They can be defined as the structural-material state unpreparedness of the part of the fault zone to fail, and a feature of the stress state (area of high effective pressure). Overcoming such small barriers requires mechanical energy, and thus, the number, density and distribution of barriers should control the reality of occurrence of a strong earthquake. If there are too many of them and the jumps of compression stress and friction force are large, then the focal is "not ripe" for a mega-earthquake. Seismological data show that large regions of a future focal may poorly manifest themselves in seismicity, and there may be a lull zone. The very size of the low stress region determines the actual magnitude of mega-earthquakes.

Thus, in the aggregate MSF corresponds to: (1) the presence of large-scale inhomogeneity of regional stress along a large fault zone; (2) the location of large area (more 300 km) low or middle stresses—the future foci of mega-earthquakes; (3) the location of a high gradient stress zone inside or near the future focal boundary of a mega-earthquake. The more pronounced the inhomogeneity of the local stress level the longer the MSF fault zone will persist. At the moment when the amplitude of this inhomogeneity is reduced to a certain minimum, there is a transition to an unstable state—a mega-earthquake.

The presence of a region of MSF precedes the emergence of mega-earthquakes. A set of accumulated data allows us to consider the moment of creation of a large region of low effective pressure as the first phase of mega-earthquake preparation, which is the longest stage in the focal preparation for a mega-earthquake. The time of zone of nucleation formation represents the final stage of preparing the mega-earthquake. For this zone, the disappearance of local stress inhomogeneity (no weak barriers) is critical. In this case, any event that occurs in the area of a local scale high and quasi-uniform stress gradient can generate a mega-earthquake.

It can be said that the area of low effective pressure, which determines the earthquake occurrence scale, is the first necessary condition of the MSF. Let's call it the *regional MSF criterion.* In fact, seismic gaps (Fedotov and Solomatin 2017) can claim this role. The zone with a high stress gradient can be considered the nucleation of a mega-earthquake. There the foreshock activity should be observed, and it can be called *a local MSF criterion.* In this zone the trigger earthquake (foreshock) should be detected. The simultaneous presence and proximity of these two zones is a signal of the final phase of MSF development and preparation for a mega-earthquake.

Thus, we can agree with the conclusion made in Brune (1979) that "the size of each earthquake is determined only in the process of rupture".

Currently, for most seismic regions, the results of tectonophysical stress inversions do not allow us to obtain the parameters of the stress field better than 15–50 km of averaging—the regional stress field. A more detailed stress distribution corresponding to the local stress field can be obtained for regions such as California, where there is a very dense network of seismic stations. The sub-regional scale stress (averaging 10–20 km) was obtained for the seismic region of the Japanese Islands (Rebetsky et al. 2016a).

4. *Discussion*

Geller et al. (1997a, b) were devoted to the analysis of earthquake prediction problems, and it was said "in order for large earthquakes to be predictable, they would have unusual events resulting from specific physical states." The specific physical state of the fault zone associated with the changes in its structure and substance is necessarily manifested in the field of tectonic stress. However, not all parameters of the stress state can equally characterize these structural-real transformations. So the orientation of the principal stress axes is more related to the external loading conditions of large rock volumes. Data on the spatial distribution of stress (deviatoric stress), effective pressure and Coulomb stress are needed to study the process of rock brittle fracture preparation and quasi-plastic flow.

Methods of estimating magnitudes of natural stress are based on a simplified form of the brittle fracture zone (strip) on a Mohr diagram and the algorithms to create homogenous sets of earthquakes with data on earthquake focal mechanisms (Rebetskii 2005; Rebetsky 2007; Rebetsky and Polets 2018), or geological slickenside set (Angelier 1989; Reches 1983; Rebetsky et al. 2017) developed today in tectonophysics.

The data on the distribution of the nature stress magnitude in earthquake foci make it possible escape from the stagnation problem of forecasting strong earthquakes. The main reason for the formation of brittle fracture in a rock mass is the heterogeneity of the stress state. This regularity is manifested in the physical experiment. Microfractures that precede the complete destruction of the sample are formed at grain boundaries and in the patches of its contact, i.e., at the scale of the maximum heterogeneity level of the rock. Accumulating and structuring strength defects create heterogeneity, manifested in the scale of the sample.

The presented tectonophysical method of allocating subduction zones as areas of future mega-earthquakes refers to long-term earthquake prediction. However, this approach allows us to note patches of earthquake trigger development or the zones of high stress gradient. These patches can be the object of observation for other methods of earthquake prediction, based on the search for indirect precursors of earthquakes in various physical fields (short-term prediction), and on the study of statistical laws of the seismic regime (medium-term prediction).

Numerical geomechanical calculations show that with homogeneous deformation at the elastic stage, a plastic flow can occur with the formation of plastic localization bands (Rice 1993). However, the solution of the brittle crack formation problem for a perfectly elastic medium in the case of a homogeneous initial stress state arising under homogeneous external loading is not correct. The solution becomes correct either if the initial state is inhomogeneous in a perfectly elastic body or if the model of the medium is an elastic–plastic body. In this case, at the stage of formation of plastic localization bands, the stress field is transformed (Rice and Uenishi 2002; Uenishi and Rice 2003; Stefanov 2005; Reches and Lockner 2010) and becomes significantly non-uniform. Brittle cracks separate the areas of the greatest deformation heterogeneity, also transforming the stress state to a more uniform one.

Currently, the possibility of a detailed study of natural stress under the conditions of the greatest homogeneity of external loading corresponds to active continental margins. There are a large number of earthquakes recorded by the global network of seismic stations IRIS, with magnitudes Mw > 4.3, for which the global CMT project massively determines the mechanisms of earthquake foci. The data on the field of natural stress in the regions of the three mega-earthquakes of the twenty-first century the Sumatra-Andaman (2004), Maule (2009), the Simushir (2006) obtained with the use of the MCA showed their significant heterogeneity in the laterals. Moreover, this heterogeneity is manifested in the field of stress magnitudes and not in the field of orientations of principal stress axes.

Since for these seismic regions stress averaging was carried out over the entire power of the conditional crust (40 km), we cannot say anything about the heterogeneity of stress in depth. The distribution of high and low stress magnitude zones (effective pressure and maximum shear stress) can be interpreted from the position of the asperity model (Lay and Kanamori 1981) and weakening of the strength (Dieterich 1992). The nature of the relationship between the asperity and areas of low stress in these three seismic regions is different.

For the South American subduction zone, the lateral asperity length is the largest from 600 to 800 km. They are separated by relatively narrow areas (50–100 km) of low stress. This ratio is interrupted in the crust of southern Chile, where the region of reduced stress has a length of approximately 700 km. There was the Maule earthquake in 2010. For the Western Flank of the Sunda Arc, the length of the asperity has the same order, but the areas of low stress separating them are somewhat wider, more than 200 km. At the northern end of one of these asperities there is an abnormally large area of low stress (approximately 600 km). This region corresponds to the focal portion of the SAE, where it allocated more than 90% of its seismic energy.

It should be noted that the results of tectonophysical zoning of the earth's crust by the intensity of the stress state along Sumatra Island coincide well with the data on the coupling distribution heterogeneity presented in Chlieh et al. (2008). Analysis of palaeogeodetic data from Chlieh et al. (2008) was carried out on the basis of the rate of coral growth subsidence and uplift.

The situation is completely different for the Kuril-Kamchatka seismic region. Here the area of asperities are a relatively few and long-range (50–100 km). These patches of high stress are separated from each other by 100–150 km sections of medium or low stress. Against this background, in the Middle and Southern Kuriles, there are two sections with lengths of approximately 500 and 250 km, respectively, within which there are no high stress patches. This is the region of foci for the Simushir (2006) and Shikotan (1994) earthquakes.

Thus, the boundaries of the mega-earthquake foci are clearly marked in the stress field as zones of increased effective pressure. Here we find the answer to the question formulated in Kagan (1997a) about the need for a deterministic description of the end of a seismogenic fault to create a forecasting system. Above, we also noted that it is a nucleation zone (high stress), which in the case of its location inside the mega-earthquake focal, can be considered a barrier; its location on the border focal can be considered the extreme part of the asperity.

Thus, the areas of future mega-earthquake foci have an individual feature in the field of tectonic stress. Here it is important to repeat the replica from the previous section. Coulomb stress in high stress zones is less than for medium and even low stress (Fig. 1). How is this difference in the regional stress states (scale averaging of stress near 50 km) of these areas manifested?

The fact is that with high levels of stress and effective pressure, brittle fracture becomes less effective. Most of the energy released in the event of failure is used to overcome frictional forces in such difficult areas and creep movement. For these regions there are few earthquakes with large magnitudes of Mw = 6.5–7.5. Earthquakes with Mw = 6.5–7.5 should occur most frequently in the regions of average stress level, against the background of a large number of events with magnitudes Mw = 4.5–6.0, which could be seismic creep. In regions of low stress, where strong earthquakes form, earthquakes with Mw $>$ 6.5 are absent. Here weak earthquakes with 3.0 $<$ Mw $<$ 5.0 are very few, i.e., it is a calm seismic zone. Once again, in all three cases we are talking about regions of the crust where the Coulomb stress reaches the limit of brittle strength.

Thus, the fact of stress state heterogeneity is critical for the development of large-scale brittle fracture along the fault. In the case of stress state quasi-homogeneity, aseismic creep or multiple acts of seismic events of medium and small magnitude (seismic creep) are more likely. The role of inhomogeneity of the stress state along the strike of shear zones was studied in Lui and Rice (2005). The reason for the different nature of the asperity distribution for different subduction zones can be considered features of the fluid regime, which is also associated with metamorphic transformations of tectonites in the fault zones of the earth's crust (Chikov 2011).

The proposed model of MSF was obtained on the basis of the analysis of nature stress states before mega-earthquakes. This model largely explains the reasons for the failure of prediction methods using statistical analysis or earthquake precursors. In particular, statistical methods of forecasting rely on the analysis of patterns of foreshocks that occur before the earthquake (the months and first years), and templates for the formation of zones of seismic calm. At the same time, the states of large regions of the crust that are the supposed foci of strong earthquakes are analysed. The task of searching and analysing the state of a small nucleation zone of a strong earthquake is never set. Therefore, a false alarm in the forecast is associated with the unpreparedness of this patch to be the trigger of such a seismic event. Skipping the forecast can be associated with a long phase of calm. Since the catalogue of mechanisms for the entire observation period (usually more than 20–30 years) is used for tectonophysical stress inversions it is possible to obtain stress data for this area. For the statistical forecast, data on earthquakes in the period 1–5 years before the predicted event are important. Their absence creates a situation in the forecast to skip events.

Another reason for the errors of prediction methods is it is known that earthquake precursors are not everywhere in the focal of a future earthquake. They arise in a limited number of fairly local areas. We assume that these may be zones of local critical states in the fault region. At the same time, the focus and its nucleation zone are not yet ready for the formation of a strong earthquake. These are false alarms in the forecast.

Thus, the obtained data on the peculiarities of the natural stress field distribution in regions of mega-earthquake show that predicting the earthquake place and magnitude differs from the forecast of time by the scale of averaged stress of study areas. Having solved the first part of the problem and singled out a dangerous area for the formation of a mega-earthquake focal, it is necessary to identify possible nucleation zones. In some cases, part of these fracture sites are visible in the field of tectonic stress on the same scale of reconstruction as the focal itself. However, in fact, monitoring of such sites requires a more detailed database of earthquake foci mechanisms with magnitudes in the range of 2.5 $<$ Mw $<$ 4.5

5. *Conclusion*

The results of tectonophysical studies of the natural stress state of active continental margin seismic regions, obtained from data on the mechanisms of earthquake foci in the range of Mw = 4.3–6.5, revealed a number of regularities. In particular, it was found that the focal of a mega-earthquake is specifically determined in the subduction area. Foci lie in areas of low level effective pressure and deviatory stress. The source of such data was the stress states that existed before the 2004 Sumatra-Andaman, the 2006 Simushir and the 2010 Maule mega-earthquakes.

It is shown that one should distinguish between the state of the main part of the mega-earthquake focal and its nucleation patch. In the main part of the focal there may be a weak variation of the effective pressure within the medium and even low levels (weak barriers). The nucleation zone is located in the patch of increased stress and effective pressure and can be located both inside the mega-earthquake focus and on its border. In this case, the propagation of the seismic rupture goes in the direction of the maximum stress gradient from the region of high to low magnitude. The results of stress analysis on the foci of the strongest twenty-first century earthquakes allowed us to give the term "metastable state" to faults, and it came to seismology from the physics of phase states and means a specific pattern of stress magnitude distribution before strong earthquakes. The proposed model of MSF answers the question from which this article began (Kagan 1997a).

In formulating a geomechanical definition of MSF in application to the problem of the mega-earthquake focus, the first task is to identify the seismic focal zone extended region (more than 200–300 km) and the small variable and low level of effective pressure. The period of metastability of this fault zone section is determined by the preparedness for the formation of a trigger earthquake in the nucleation zone with Mw = 6.5–7.0. The duration of this period may be the years or tens years. The metastable stage is preceded by a stable state for which the extended region of the reduced level of effective stress has large variation in magnitude. The transition from MSF to unstable is determined by processes in the nucleation zone and requires additional research, both with laboratory experiment and using tectonophysical stress inversion for earthquakes of Mw = 2.5–5.5. Allocation of the stress gradient patch distribution determines the epicentre of mega-earthquake and makes it possible to implement comprehensive studies of large seismogenic structures to identify the phase of "maturation" for strong earthquakes.

Acknowledgements

Funding was provided by Russian Foundation for Fundamental Investigations (Grant nos. 16-05-01115 and 17-05-01193), by the program of the National Natural Science Foundation of China (Grant No. 41572181) and by the state program on the IPE RAS.

REFERENCES

Abercrombie, R. E., & Rice, J. R. (2005). Can observations of earthquake scaling constrain slip weakening? *Geophysical Journal International, 162,* 406–424. https://doi.org/10.1111/j.1365-246X.2005.02579.x.

Aki, K. (1984). Asperities, barriers, characteristic earthquakes and strong motion predictio. *Journal of Geophysical Research, 89*(B7), 5867–5872.

Ammon, C. J., Ji, C., Thio, H.-K., et al. (2005). Rupture process of the 2004 Sumatra-Andaman earthquake. *Science*, *308*, 1133–1139.

Anderson, E. M. (1951). The dynamics of faulting. *Transactions of Edinburgh Geological Society, 8,* 387–402.

Angelier, J. (1979). Determination of mean principal directions of stresses for a given fault population. *Tectonophysics, 56,* T17–T26.

Angelier, J. (1989). From orientation to magnitude in paleostress determinations using fault slip data. *Journal of Structural Geology, 11*(1/2), 37–49.

Angelier, J. (1990). Inversion field data in fault tectonics to obtain the regional stress—III. A new rapid direct inversion method by analytical means. *Geophysical Journal International, 10,* 363–367.

Bakun, W. H., Aagaard, B., Dost, B., Ellsworth, W. L., Hardebeck, J. L., Harris, R. A., et al. (2005). Implications for prediction and hazard assessment from the 2004 Parkfield earthquake. *Nature, 437,* 969–974.

Brace, W. F., & Byerlee, J. D. (1966). Stick slip as a mechanism for earthquakes. *Science, 153,* 990–992.

Brune, J. N. (1979). Implications of earthquake triggering and rupture propagation for earthquake prediction based on premonitory phenomena. *Journal of Geophysical Research, 84,* 2195–2198.

Byerlee, J. D. (1970). Static and kinetic friction of granite under high stress. *International Journal of Rock Mechanics and Mining Sciences, 7,* 577–582.

Byerlee, J. D. (1978). Friction of rock. *Pure and Applied Geophysics, 116,* 615–626.

Chikov, B. M. (2011). *Introduction to the physical basis of static and dynamic geomechanics* (p. 299). Novosibirsk: GEO **(in Russ)**.

Chlieh, M., Avouac, J. P., Sieh, K., Natawidjaja, D. H., & Galetzka, J. (2008). Heterogeneous coupling of the Sumatran megathrust constrained by geodetic and paleogeodetic measurements. *Journal of Geophysical Research, 113,* L21603. https://doi.org/10.1029/2007JB004981.

Cicerone, R. D., Ebel, J. E., & Britton, J. (2009). A systematic compilation of earthquake precursors. *Tectonophysics, 476*(3), 371–396.

Corbi, F., Funiciello, F., Brizzi, S., Lallemand, S., & Rosenau, M. (2017). Control of asperities size and spacing on seismic behavior of subduction megathrusts. *Research Letters*. https://doi.org/10.1002/2017gl074182.

Das, S., & Aki, K. (1977). Fault planes with barriers; a versatile earthquake model. *Journal of Geophysical Research, 82,* 5658–5670.

Davis, C., Keilis-Borok, V., Kossobokov, V., & Soloviev, A. (2012). Advance prediction of the March 11, 2011 Great East Japan Earthquake: A missed opportunity for disaster preparedness. *International Journal of Disaster Risk Reduction, 1,* 17–32. https://doi.org/10.1016/j.ijdrr.2012.03.001.

Dieterich, J. (1972). Time-dependence of rock friction. *Journal of Geophysical Research, 77,* 3690–3697.

Dieterich, J. H. (1979). Modeling of rock friction—1 experimental results and constitutive equations. *Journal of Geophysical Research, 84,* 2161–2168.

Dieterich, J. H. (1992). Earthquake nucleation on faults with rate- and state-dependent strength. *Tectonophysics, 211*(1–4), 115–134.

Fedotov, S. A. (1965). Regularities of the distribution of strong earthquakes in Kamchatka, the Kurile Islands and northeastern Japan (in Russian). *Tr Inst Fiz Zemli Akad Nauk SSSR, 36*(203), 66–93.

Fedotov, S. A. (1968). *On the seismic cycle, the possibility of quantitative zoning and long-term seismic prediction Book Seismic zoning of the USSR* (pp. 121–150). Moscow: Nauka. **(in Russ)**.

Fedotov, S. A., & Solomatin, A. V. (2017). Long-term seismic forecast for the Kuril-Kamchatka island arc for the April 2016 though March 2021, its development and application; characteristics of seismic activity of the Kuril-Kamchatka arc before and after the deep sea of Okhotsk earthquake of 24.V 2013, M = 8.3. *Volcanology and Seismology, 11*(3), 173–186.

Frankel, A. (2013). Comment on ‘‘Why earthquake hazard maps often fail and what to do about it". *Tectonophysics, 592,* 200–206.

Geller, R. J. (1997). Earthquake prediction: A critical review. *Geophysical Journal International, 131,* 425–450.

Geller, R. J., Jackson, D. D., Kagan, Y. Y., & Mulargia, F. (1997a). Response—cannot earthquakes be predicted? *Science, 278,* 488–490.

Geller, R. J., Jackson, D. D., Kagan, Y. Y., & Mulargia, F. (1997b). Earthquakes cannot be predicted. *Science, 275,* 1616–1617.

Geller, R. J., Mulargia, F., & Stark, P. B. (2015). Why we need a new paradigm of earthquake occurrence. In G. Morra, et al. (Eds.), *Subduction dynamics: From mantle flow to mega disasters, geophysical monograph* (Vol. 211, pp. 183–191). Washington, DC: American Geophysical Union.

Gilbert, G. K. (1884). A theory of the earthquakes of the Great Basin, with a practical application. *American Journal of Science, Series 3, 27*(157), 49–54.

Gol'din, S. V. (2004). Dilatancy, repacking and earthquakes. Izvestiya. *Physics of the Solid Earth, 40*(10), 817–832.

Gol'din, S. V. (2005). Macro and mesostructure focal region. *Fiz Zemli, 41*(1), 1–5. **(in Russ)**.

Gushchenko, O. I. (1975). Kinematic principle of reconstruction of main stress directions (according to geological and seismological data). *Doklady AN USSR Series Geophysics, 225*(3), 557–560. **(in Russ)**.

Gushchenko, O. I. (1996). Seismotectonic stress monitoring of the lithosphere (structural and kinematic principle and the main elements of the algorithm). *Doklady RAN, 346*(3), 399–402. **(in Russ)**.

Gzovsky, M. V. (1954a). On the problems and content of Tectonophysics. *Izvestiya USSR Academy of Sciences of the Geophysics Series, 3,* 244–263. **(in Russ)**.

Gzovsky, M. V. (1954b). Tectonic stress fields. *Izvestiya USSR Academy of Sciences of the Geophysics Series, 3,* 390–410. **in Russ**.

Gzovsky, M. V. (1957). Tectonophysics substantiation of geological criteria of seismicit (I and II). *Izvestiya USSR Academy of Sciences of the Geophysics Series, 141*(2), 273. **(in Russ)**.

Gzovsky, M. V. (1975). *Fundamentals of tectonophysics*. Moscow: Nauka. **(in Russ)**.

Jackson, D. D., & Kagan, Y. Y. (2006). The 2004 Parkfield earthquake, the 1985 prediction, and characteristic earthquakes: Lessons for the future. *Bulletin of the Seismological Society of America, 96,* S397–S409.

Jackson, D. D., & Kagan, Y. Y. (2011). Characteristic earthquakes and seismic gaps. In H. K. Gupta (Ed.), *Encyclopedia of solid earth geophysics* (pp. 37–40). Berlin: Springer. https://doi.org/10.1007/978-90-481-8702-7.

Jordan, T. H. (2006). Earthquake predictability, brick by brick. *Seismological Research Letters, 77*(1), 3–6.

Kagan, Y. Y. (1997a). Are earthquakes predictable. *Geophysical Journal, 131,* 505–525.

Kagan, Y. Y. (1997b). Statistical aspects of Parkfield earthquake sequence and Parkfield prediction experiment. *Tectonophysics, 270,* 207–219.

Kagan, Y. Y., & Jackson, D. D. (1991). Seismic gap hypothesis: Ten years after. *Journal of Geophysical Research, 96,* 21419–21431.

Kagan, Y. Y., & Jackson, D. D. (2013). Tohoku earthquake: A surprise? *Bulletin of the Seismological Society of America, 103,* 1181–1194.

Kagan, Y. Y., Jackson, D. D., & Geller, R. J. (2012). Characteristic earthquake model, 1884–2011, R.I.P. *Seismological Research Letters, 83,* 951–953.

Kanamori, H. (1978). Nature of seismic gaps and foreshocks. In *Proceedings of conference VI: Methodology for identifying seismic gaps and soon to break gaps, USGS Open File Report 78-943* (pp. 319–334).

Kanamori, H., & Stewart, G. S. (1978). Seismological spectsof the Guatemala earthquake of February 4, 1976. *Journal of Geophysical Research, 83,* 3427–3434.

Karato, Sh. (2008). *Deformation of earth materials. An introduction to the rheology of solid earth*. Cambridge: Cambridge University Press.

Keilis-Borok, V. I., & Kossobokov, V. G. (1990). Premonitory activation of seismic flow: Algorithm M8. *Physics of the Earth and Planetary Interiors, 61,* 73–83.

Kenzie, Mc, & Dan, P. (1969). The relation between fault plane solutions for earthquakes and directions of the principal stresses. *Bulletin of the Seismological Society of America, 59*(2), 591–601.

Kocharyan, G. G., & Batukhtin, V. I. (2018). Laboratory studies of the fault sliding process as a physical basis for a new approach to short-term earthquake prediction. *Geodynamics and Tectonophysics, 9*(3), 671–691. **(in Russ)**.

Kocharyan, G. G., & Novikov, V. A. (2016). Experimental study of different modes of block sliding along interface. Part 1. Laboratory experiments. *Physical Mesomechanics, 19*(2), 189–199. https://doi.org/10.1134/S1029959916020120.

Kossobokov, V. G., Healy, J. H., & Dewey, J. W. (1997). Testing an earthquake prediction algorithm. *PAAG, 149,* 219–232.

Kostrov, B. V. (1975). *Mechanics of tectonic earthquake focal*. Moscow: Nauka. **(in Russ)**.

Langbein, J., Borchrdt, R., Dreger, D., Fletcher, J., Hardebeck, J. L., Hellweg, M., et al. (2005). Preliminary report on the 28 September 2004 M 6.0 Parkfield. *California earthquake. Seismological Research Letters, 76*(1), 10–26.

Lay, T., & Kanamori, H. (1980). Earthquake doublets in the Solomon Islands. *Physics of the Earth and Planetary Interiors, 21,* 283–304.

Lay, T., & Kanamori, H. (1981). An asperity model of great earthquake sequences. In D. W. Simpson & P. G. Richards (Eds.), *Earthquake prediction—an international review, Maurice Ewin Series* (Vol. 4, pp. 579–592). Washington DC: American Geophysical Union.

Lay, T., Kanomory, H., Ammon, C. J., et al. (2005). The great Sumatra-Andaman earthquake of 26 December 2004. *Science, 308,* 1127–1133.

Liu, Y., & Rice, J. R. (2005). Aseismic slip transients emerge spontaneously in three-dimensional rate and state modeling of subduction earthquake sequences. *Journal of Geophysical Research, 110,* B08307. https://doi.org/10.1029/2004JB003424.

Ma, J., & Guo, Y.-S. (2014). Accelerated synergism prior to fault in stability: Evidence from laboratory experiments and an earthquake case (in Chinese). *Seismology and Geology, 36*(3), 548–562. https://doi.org/10.3969/j.issn.0253-4967.2014.03.001.

Ma, J., Sherman, S. I., & Guo, Y.-S. (2012). Identification of meta-instable stress state based on experimental study of evolution of the temperature field during stick-slip instability on a 5° bending fault. *Science China Earth Sciences, 55*(6), 869–881. https://doi.org/10.1007/s11430-012-4423-2.

Mikumo, T., & Miyatake, T. (1978). Dynamic rupture process on a threedimensional fault with non-uniform frictions and near-field seismic waves. *Geophysical Journal of the Royal Astronomical Society, 54,* 417–438.

Mogi, K. (1962). Study of elastic shocks caused by the fracture of heterogeneous material and its relation to earthquake phenomena. *Earthquake Research Institute, The University of Tokyo, 40,* 125–173.

Mori, J., Abercrombie, R. E., & Kanamori, H. (2003). Stress drops and radiated energies of aftershocks of the 1994 Northridge, California, earthquake. *Journal of Geophysical Research, 108*(B11), 2545. https://doi.org/10.1029/2001JB000474.

Mulargia, F. (2013). Why the next large earthquake is likely to be a big surprise. *Bulletin of the Seismological Society of America, 103,* 2946–2952.

Mulargia, F., & Bizzarri, A. (2015). Fluid pressure waves trigger earthquakes. *Geophysical Journal International, 200,* 1279–1283.

Nishenko, S. P., & Sykes, L. R. (1993). Comment on ‘‘Seismic gap hypothesis: ten years after” by Y.Y. Kagan and D.D. Jackson. *Journal of Geophysical Research, 98,* 9909–9916.

Osokina, D. N. (1987). *The relationship of displacements along faults with tectonic stress fields and some issues of the destruction of the mountain range. Book: The stress field and deformations in the earth's crust* (pp. 120–135). Moscow: Nauka. **(in Russ)**.

Panza, G., Kossobokov, V. G., Peresan, A., & Nekrasova, A. (2014). Why are the standard probabilistic methods of estimating seismic hazard and risks too often wrong. In M. Wyss & J. F. Shroder (Eds.), *Earthquake hazard, risk and disasters* (pp. 309–357). Amsterdam: Elsevier.

Rebetskii, Y. L. (1997). Reconstruction of tectonic stresses and seismotectonic strain: Methodical fundamentals, current stress field of Southeastern Asia and Oceania. *Translation (Doklady) of the Russian Academy of Science, 354*(4), 560–563.

Rebetskii, Y. L. (2003). Development of the method of cataclastic analysis of shear fractures for tectonic stress estimation. *Doklady Earth Sciences, 388*(1), 72–76.

Rebetskii, Y. L. (2005). Tectonic stress, metamorphism, and earthquake source model. *Doklady Earth Sciences, 400*(1), 127–131.

Rebetsky, Y. L. (1996). I. Stress-monitoring: Issues of reconstruction methods of tectonic stresses and seismotectonic deformations. *Journal of Earthquake Prediction Research. Beijing, China, 5*(4), 557–573.

Rebetsky, Y. L. (2006). *Dilatancy, pore fluid pressure and new data on the strength of rock massifs in natural bedding. Book: Fluid and geodynamics* (pp. 120–146). Moscow: Nauka. **(in Russ)**.

Rebetsky, Y. L. (2007). *Tectonic stresses and strength of rock massifs*. Moscow: Akademkniga. **(in Russ)**.

Rebetsky, Yu L. (2009). Stress state of the Earth's crust of the Kurils Islands and Kamchatka before the Simushir earthquake. *Russian Journal of Pacific Geology, 3*(5), 477–490. https://doi.org/10.1134/S1819714009050108.

Rebetsky, Yu L. (2015). On the specific state of crustal stresses in intracontinental orogens. *Geodynamics and Tectonophysics, 6*(4), 437–466. https://doi.org/10.5800/GT-2015-6-4-0189.

Rebetsky, Y. L., & Marinin, A. V. (2006). Preseismic stress field before Sumatra-Andaman earthquake of 26.12.2004. A model of metastable state of rocks. *Russian Geology and Geophysics, 47*(11), 1173–1185.

Rebetsky, Y. L., & Polets, A. Y. (2018). The method of cataclastic analysis of discontinuous displacements. Moment tensor solutions—a useful tool for seismotectonics. In S. D'Amico (Ed.), *Moment tensor solutions. Springer natural hazards* (pp. 112–162). Cham: Springer. https://doi.org/10.1007/978-3-319-77359-9_6.

Rebetsky, Yu L, Polets, A Yu., & Zlobin, T. K. (2016a). The state of stress in the Earth's crust along the northwestern flank of the Pacific seismic focal zone before the Tohoku earthquake of 11 March 2011. *Tectonophysics, 685,* 60–76. https://doi.org/10.1016/j.tecto.2016.07.016.

Rebetsky, Y. L., Sim, L. A., & Marinin, A. V. (2017). *From slickensides to tectonic stresses. Methods and algorithms*. Moscow: GEOS. **(in Russ)**.

Rebetsky, Yu L, Sycheva, N. A., Kuchay, O. A., & Tatevossian, R. E. (2012). Development of inversion methods on fault slip data.

Stress state in orogenes of the central Asia. *Tectonophysics, 581,* 114–131. https://doi.org/10.1016/j.tecto.2012.09.027.

Rebetsky, Yu L, Sycheva, N. A., Sychev, V. N., Kuzikov, S. I., & Marinin, A. V. (2016b). The stress state of the northern Tien Shan crust based on the KNET seismic network data. *Russian Geology and Geophysics, 57,* 387–408. https://doi.org/10.1016/j.rgg.2016.03.003.

Rebetsky, Yu L, & Tatevossian, R. E. (2013). Rupture propagation in strong earthquake sources and tectonic stress field. *Bulletin de la Societe Geologique de France, 184*(4–5), 335–346.

Reches, Z. (1983). Faulting of rock in three dimensional strain fields. II Theoretical analysis. *Tectonophysics, 95,* 133–156.

Reches, Z., & Lockner, D. A. (2010). Fault weakening and earthquake instability by powder lubrication. *Nature, 467,* 452–455. https://doi.org/10.1038/nature09348.

Reid, H. F. (1911). The elastic-rebound theory of earthquakes. *University of California Publications. Bulletin of the Department of Geological Sciences, 6,* 413–444.

Rice, J. (1982). The mechanics of earthquake rupture. In A. Dziewonski & E. Boschi (Eds.), *Physics of the earth's interior* (pp. 555–649). Amsterdam: Elsevier.

Rice, J. R. (1993). Spatio-temporal complexity of slip on a fault. *Journal of Geophysical Research, 98,* 9885–9907.

Rice, J. R. (2000). Fracture energy of earthquakes and slip-weakening rupture parameters. In *EOS transactions American geophysical union, fall meeting supplement* (Vol. 81, No. 48, p. F1227).

Rice, J. R., Uenishi, K. (2002). Slip development and instability on a heterogeneously loaded fault with power-law slip-weakening. In *EOS transactions American geophysical union, fall meeting supplement* (Vol. 83, No. 47) **(abstract S61E-06)**.

Rubin, A. M., & Ampuero, J.-P. (2005). Earthquake nucleation on (aging) rate and state faults. *Journal of Geophysical Research, 110*(B1), 1312. https://doi.org/10.1029/2005JB003686.

Ruina, A. (1983). Slip instability and state variable friction laws. *Journal of Geophysical Research, 88,* 10359–10370.

Scholz, C. H., & Engelder, T. (1976). Role of asperity indentation and ploughing in rock friction. *International Journal of Rock Mechanics and Mining Sciences, 13,* 149–154.

Shimazaki, K., & Nakata, T. (1980). Time-predictable recurrence model for large earthquakes. *Geophysical Research Letters, 7,* 279–282.

Sieh, K. E. (1978). Prehistoric large earthquakes prodused by slip on the San Sndreas fault at Pallett Creek, California. *Journal of Geophysical Research, 83*(B8), 3907–3939.

Sobolev, G. A. (2011). Seismicity dynamics and earthquake predictability. *Natural Hazards and Earth Systems Sciences, 11,* 1–14.

Sobolev, G. A., & Lyubushin, A. A. (2007). Microseismic anomalies before the earthquake on December 26, 2004 in Sumatra. *Physics of the Earth, 5,* 3–16. **(in Russ)**.

Stefanov, Yu P. (2005). Some features of the numerical modeling of elastic-large-blocks material. *Physical Mesomechanics, 8*(3), 129–142. **(in Russ)**.

Uenishi, K., & Rice, J. R. (2003). Universal nucleation length for slip weakening rupture instability under non-uniform fault loading. *Journal of Geophysical Research*. https://doi.org/10.1029/2001jb001681.

Zoback, M. L. (1992). First- and second-order patterns of stress in the lithosphere: the world stress map project. *Journal of Geophysical Research Atmospheres*, *97*(B8), 11703–11728.

(Received September 8, 2018, revised February 3, 2019, accepted February 7, 2019)