ISSN 1028-334X, Doklady Earth Sciences, 2006, Vol. 408, No. 4, pp. 635–639. © Pleiades Publishing, Inc., 2006.
Original Russian Text © V.V. Bogdanov, 2006, published in Doklady Akademii Nauk, 2006, Vol. 408, No. 3, pp. 393–397.

GEOPHYSICS

A Probabilistic Interpretation of the Law of Earthquake Recurrence: Case Study of the Kamchatka Region

V. V. Bogdanov

Presented by Academician A.S. Alekseev February 28, 2005

Received May 5, 2005

DOI: 10.1134/S1028334X06040283

It is well known that earthquake recurrence plots are well described by the exponential function $N(k) = A_0 \cdot 10^{-\gamma(k-k_0)}$, where A_0 is the seismic activity reduced to surface and time units; $k = \log E$ is the energy class; and γ is the angle factor. The relation $\log N(k) = \log A_0 - \gamma(k - k_0)$ on a double-logarithmic scale appears as straight lines. However, for the given time interval, A_0 is a function of coordinates and, therefore, a variable value even in a single seismoactive region but in different unit areas. Therefore, the arising challenge is to find such a representation of the law of recurrence (LR), in order to make it statistically invariable in terms of the transformation of coordinates. Such an opportunity appears in the case of expansion of theoretic-probabilistic concepts on event catalogues of random seismoactive regions. The present paper is devoted to the solution of this problem on the basis of a Kamchatka catalogue case study and the discussion of certain relationships revealed by the analysis of seismic regimes based on the proposed method.

In the theoretic-probabilistic approach, an earthquake is examined as elementary event ω_i in elementary event space Ω [2]. Each single event ω_i is characterized by a system of continuous stochastic variables (energy class k, latitude φ, longitude λ, depth h, and time t). In this paper, time is excluded from the system of stochastic variables. The seismicity of either the entire region or a selected part thereof is regarded as a complete group of events and described as distributions of conditional and unconditional probabilities P represented in frequency form. Stochastic events are defined as combinations of the system of stochastic variables k, φ, λ, and h within the set $\tilde{F}$. Thus, one can present the catalog of seismic events for the observation period as a stochastic space of three objects $\{\Omega, \tilde{F}, P\}$ and calculate probability distributions for various stochastic events. If the law of distribution of a system of stochastic variables is analytically specified by the distribution function $F(\varphi, \lambda, h, k)$ or by its density $f(\varphi, \lambda, h, k)$, then distribution patterns for individual variables can be found from standard formulas. In our formulation, an inverse representation of the problem will be the most logical one: deduction of the law of system distribution from the distribution patterns of stochastic variables. The density of distribution for continuous variables can be represented as either the 4th-order mixed partial derivative of $F(\varphi, \lambda, h, k)$ or the product of conditional and unconditional f functions by the following relation:

$$f(\varphi, \lambda, h, k) = \frac{\partial^4 F(\varphi, \lambda, h, k)}{\partial\varphi\, \partial\lambda\, \partial h\, \partial k}$$
$$= f(\varphi) f(\lambda|\varphi) f(h|\varphi, \lambda) f(k|\lambda, \varphi, h), \qquad (1)$$

where $f(\varphi)$ is the unconditional density of the distribution of events as a function of φ; $f(\lambda|\varphi)$ is the density of an event's distribution along λ, provided the latitude is φ; $f(h|\varphi, \lambda)$ is the distribution density along h, provided the latitude and longitude are φ and λ, respectively; $f(k|\lambda, \varphi, h)$ is the density of distribution along k, provided the longitude, latitude, and depth are λ, φ, and h, respectively. Knowing the analytical formula (1), it is possible to compute the probability of an event falling within the given intervals of latitude $\Delta\varphi_i$, longitude $\Delta\lambda_j$, depth Δh_m, and class Δk_n:

$$P(\Delta\varphi_i, \Delta\lambda_j, \Delta h_m, \Delta k_n)$$
$$= \int_{\varphi_1}^{\varphi_2} d\varphi \int_{\lambda_1}^{\lambda_2} d\lambda \int_{h_1}^{h_2} dh \int_{k_1}^{k_2} f(k, \varphi, \lambda, h) dk$$
$$= F(\varphi_i, \lambda_j, h_m, k_n) - F(\varphi_{i-1}, \lambda_{j-1}, h_{m-1}, k_{n-1})$$

*Institute of Cosmophysical Research and Radio Wave Propagation, Far East Division, Russian Academy of Sciences, ul. Mirnaya 7, selo Paratunka, Elizovskii raion, Kamchatka oblast, 684034 Russia;
e-mail: vbogd@ikir.kamchatka.ru*

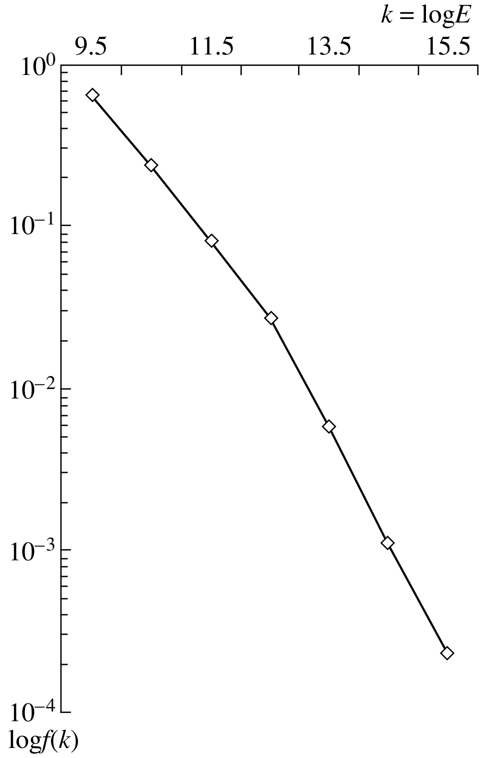

Fig. 1. A smoothed double-logarithmic plot of distribution density $f(k)$. Marked along the x-axis are mid-intervals ($k_i \pm 0.5$). The total number of seismic events for area S_1 of the raw earthquake catalogue $N_t = 21845$.

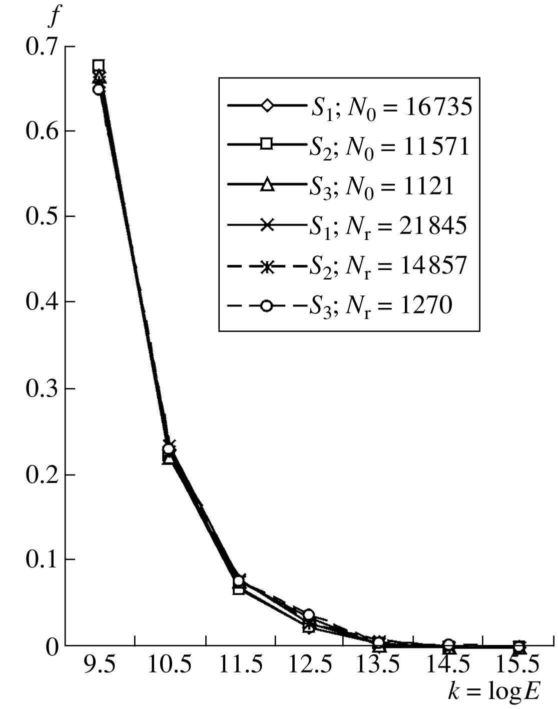

Fig. 2. Distribution density $f(k)$ of the earthquake RP of the aftershock-free and raw earthquake catalogues with hypocenters located at areas S_1 (50°–59° N, 153°–168° E), S_2 (51°–56° N, 156°–163° E), and S_3 (52°–53° N, 159°–160° E). (N_0) the number of events in the aftershock-free catalogue; (N_r) the number of events in the raw catalogue; ($k_0 = 19.5$) the initial value of energy class.

$$= P(\Delta\varphi_i)P(\Delta\lambda_i|\Delta\varphi_i)$$
$$\times P(\Delta h_m|\Delta\lambda_j, \Delta\varphi_i)P(\Delta k_n|\Delta h_m, \Delta\lambda_j, \Delta\varphi_i), \quad (2)$$

where i, j, m, and n are indices of intervals of the respective stochastic variables. Numerical values of $P(\Delta k_i, \Delta\varphi_j, \Delta\lambda_m, \Delta h_n)$ based on the seismic event frequency data catalogue are easily computable. One can similarly calculate unconditional distribution patterns for all the stochastic variables k, φ, λ, and h, as well as various combinations for conditional distribution patterns of this variables. Catalogue processing based on Eq. (2) makes it possible to compute frequencies of seismic event occurrence within a given interval of stochastic variables k, φ, λ, h and to obtain values of the distribution function $F(\Delta\varphi, \Delta\lambda, \Delta h, \Delta k)$.

On the basis of a raw catalogue of Kamchatka earthquakes and Eq. (2), we plotted a bar chart of seismic events within the energy class intervals $k_i \pm 0.5$ and $\Delta k = 1$ beginning with the representative class $k_0 = 9.5$ for the January 1, 1962 to December 31, 1999 period ($T = 37$ yr). Epicenters of the events ($N_r = 21845$) fall at the area S_1 ($\Delta\varphi = 50°–59°$ N, $\Delta\lambda = 153°–168°$ E). The smoothed bar chart is approximated by the formula:

$$f(k^{1,2}) = f(k_0^{1,2}) \cdot 10^{-\gamma_{1,2}(k^{1,2} - k_0^{1,2})}. \quad (3)$$

On a double-logarithmic scale, Eq. (3) takes the following shape:

$$\log f(k^{1,2}) = \log f(k_0^{1,2}) - \gamma_{1,2}(k^{1,2} - k_0^{1,2}). \quad (4)$$

The relationship (4) is presented in Fig. 1. The interval 12.5 ± 0.5 shows a plot bend dividing the recurrence pattern into two segments with different plot inclination angles γ. Formulas (3) and (4) refer to the following k variation ranges: $9.5 \le k_1 < 12.5$ with $\gamma_1(k_0^1 = 9.5)$, and $12.5 \le k_2 < 16$ with γ_2 $(k_0^2 = 12.5)$. Although Eqs. (3) and (4) are similar with well-known equations based on the RP, the former equations lack the dependence on the seismic activity A_0. Since density $f(k_0)$ is the initial variable in Eqs. (3) and (4), the question arises of how $f(k_0)$ varies depending on the study area. Figure 2 shows the smoothed density distribution histograms $f(k)$ for epicenters situated at time T in areas $S_1 > S_2 > S_3$ (S_2: $\Delta\varphi = 51°–56°$ N and $\Delta\lambda = 156°–163°$ E; S_3: $\Delta\varphi = 52°–53°$ N and $\Delta\lambda = 159°–160°$ E). The histograms correspond to the raw (dashed lines) and aftershock-free catalogues with initial $k_0 = 9.5$. The earthquake data were kindly placed at our disposal by the Kamchatka Branch, Geophysical Service, Russian Academy of Sciences. This catalogue for the period T with epicenters located in area S_1 was cleared of aftershocks of large Kamchatka events (algorithm compiled by G.M. Molchan and O.E. Dmitrieva; program compiled by V.B. Smirnova; data provided by V.A. Saltykov). For $\Delta k = 1$, the probabilities $P(k) = f(k)\Delta k$ are numerically the same as $f(k_0)$. Analysis of Fig. 2 reveals an interesting stability of the respective frequencies. The given minimum k values show a virtual coincidence of the plots, irrespective of the involved area and the earthquake catalogue type (aftershock-free or raw). Insignificant differences are noted for the recurrence plot with the least area S_3, for which the number of seismic events is one order of magnitude less than for areas S_1 and S_2. The higher is the number of involved events N, the better the coincidence. This conclusion is consistent with limiting theorems of probability theory. Therefore, it is possible to suggest a statistical description of the earthquake RP in probabilistic representation based on the invariant Eqs. (3) and (4), irrespective of the area involved. This is an important distinction as compared with the recurrence plots, with their representation based on the concept of seismic activity A_0 [1], [3]. Under such an approach, the recurrence pattern acquires a probabilistic interpretation, does not depend on the study area, and characterizes features of the medium during its destruction.

Figure 3 shows averaged double-logarithmic plots of the density distribution $f(k)$ for the aftershock-free and raw catalogues ($k_0 = 9.5$) drawn on the basis of Fig. 2. For comparison sake, we also present similar plots for $k_0 = 10.5$ and average $f(k)$ values with the initial k_0 equal to 8.5 (areas S_3, S_4: $\Delta\varphi = 52°–52.5°$ N and $\Delta\lambda = 159°–159.5°$ E) and 11.5 (areas S_1, S_2, and S_3), respectively. Quite noticeable is the practical coincidence of $f(k)$ plots for the aftershock-free and raw catalogues under the same k_0 values equal to 9.5 and 10.5, respectively. The calculations show that the average probabilities for various initial k_0 values based on processing of the aftershock-free and raw earthquake catalogues for $S_1 > S_2 > S_3 > S_4$ ($\Delta k = 1$) are practically the same: P_{S_3, S_4} ($k_0 = 8.5$) = 0.648, P_{S_1, S_2, S_3} ($k_0 = 9.5$) = 0.662, P_{S_1, S_2, S_3} ($k_0 = 10.5$) = 0.674, and P_{S_1, S_2, S_3} ($k_0 = 11.5$) = 0.678 (the preference for S_3 and S_4 was determined by their correspondence to the areas of reliable registration of events with $k \geq 8.5$).

An important feature of probabilistic representation of the recurrence pattern is the integral of $f(k)$ over the interval from the initial value k_0 to k_{max} (in a mathematic sense, $k = \infty$) being equal to one. Calculating from (3) the sum of integrals of $f(k^1)$ and $f(k^2)$ over the $k_0^1 - \frac{\Delta k}{2}$ to k_0^2 and k_0^2 to ∞ intervals, respectively, and keeping in mind that $\gamma_2 < 0$, we will obtain:

$$\frac{f(k_0^1)}{\gamma_1 \ln 10} \cdot 10^{\gamma_1 \Delta k/2}[1 - 10^{-\gamma_1(k_0^2 - k_0^1 + \Delta k/2)}] - \frac{f(k_0^2)}{\gamma_2 \ln 10} = 1. \quad (5)$$

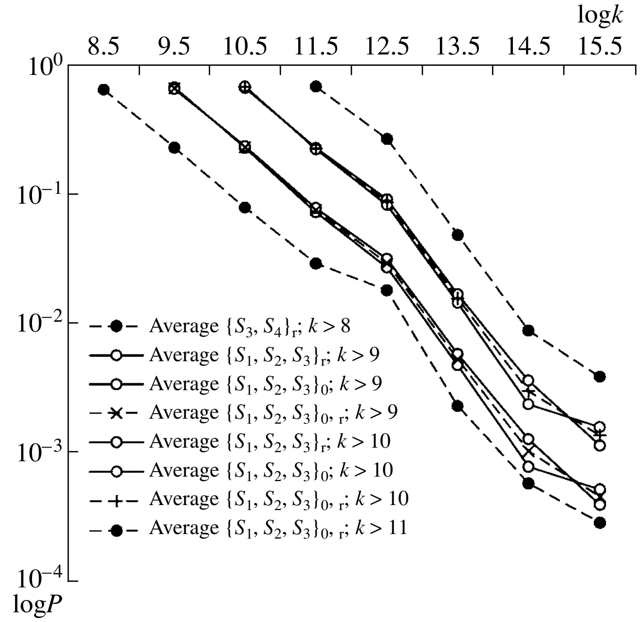

Fig. 3. The averaged for areas S_1, S_2, and S_3 double-logarithmic distribution densities $f(k)$; dashed lines show averaged distributions $f(k)$ for the aftershock-free and raw catalogues, k_0 = 8.5, 9.5, 10.5, and 11.5.

Eq. (5) demonstrates the obvious similarity of the plots shown in Fig. 3 with the practically similar computed initial values of $f(k_0^1)$. Indeed, when the distribution density $f(k)$ is integrated from the fixed value k_0 to infinity, and the integral is normalized to unity, the lower limit bias on the finite $\pm\Delta k$ will not entail any modification of the initial values of $f(k_0^1)$.

From the condition of equality of the two functions specified by (3), we obtain the second equation at the point k_0^2:

$$f(k_0^1) \cdot 10^{-\gamma_1(k_0^2 - k_0^1)} = f(k_0^2). \quad (6)$$

Expressions (5) and (6) allow us to easily compute $f(k_0^1)$ and $f(k_0^2)$ from known initial values γ_1 and γ_2. For the initial $k_0 = 9.5$, we have $f(k_0^1) = 0.662$; for $k_0^2 = 12.5$, we have $f(k_0^2) = 0.029$. From (5) and (6), we obtain $\gamma_1 = 0.453$ and $\varphi_2 = 0.679$. From the similarity of Fig. 3 plots, it follows that inclination factors γ for various initial k_0 will also be practically equal.

However, an alternative solution is possible. Having processed an earthquake catalogue for intervals ($k_i \pm 0.5$), we evaluate $f(k_i)$ and $\log f(k_i)$ and plot the respective relationships for each linear segment by the least-squares method. Figure 4 provides such relations for the initial $k_0 = 9.5$. It also gives their mathematical representations for each segment of the dashed curve

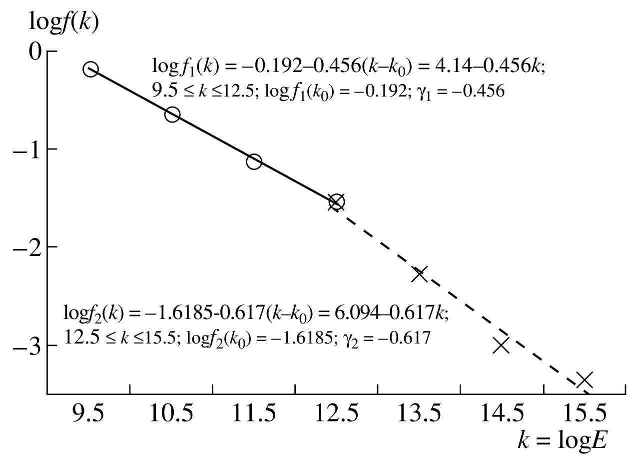

Fig. 4. Double-logarithmic analytical relations $f(k)$ based on least squares fitting of averaged distributions for the aftershock-free and raw catalogues ($k_0 = 9.5$).

$\log f(k_i)$, as well as the respective numerical values for γ_1 and γ_2. In a similar way, for the initial $k_0 = 10.5$, the relationship $\log f(k_i)$ can be plotted in the following way:

For the first segment $10.5 \le k \le 12.5$:

$$\log f_1(k) = -0.175 - 0.447(k - k_0^1)$$
$$= 4.5185 - 0.447k; \quad \log f_1(k_0^1) = -0.175;$$
$$\gamma_1 = -0.447;$$

For the second segment $12.5 \le k \le 15.5$:

$$\log f_2(k) = -1.1445 - 0.617(k - k_0^2)$$
$$= 6.568 - 0.617k;$$
$$\log f_2(k_0^2) = -1.1445; \quad \gamma_2 = -0.617.$$

Comparing the values of the inclination factor γ for respective segments of the dashed lines $\log f(k_i)$ with different initial k_0 values, we note that $\gamma_2 = 0.617$ for both relations with $k \ge 12.5$. For the relations with $k \le 12.5$, the γ values are slightly different, which can be attributed to the different numbers of points involved in the least-squares calculation of inclinations for $k_0 = 9.5$ (4 points) and $k_0 = 10.5$ (3 points). Having compared the factors γ obtained by the least-squares method for $k_0 = 9.5$ with those computed from the (5) and (6) system, we note that γ_1 values are practically equal (discrepancy 0.003), but the difference between γ_2 values amounts to 0.062. The least-squares calculation of γ_2 in the second segment of $\log f(k_i)$ is more physical, because we use the real events falling at the interval with $k_i = \max$, in contrast to the mathematical approximation with $k_{\max} \to \infty$. Precisely, this reason is responsible for the difference in γ_2 values obtained by two different methods.

Let us multiply by Δk both the left and right part of Eq. (1). Having omitted the numerical indices "1, 2" for the sake of simplification and keeping in mind that $f(k_i)\Delta k = P(k_i)$, where $f(k_i)$ and $P(k_i)$ are the i-th Δk mid-interval values of f and P, we obtain:

$$P(k_i) = P(k_0) \cdot 10^{-\gamma(k_1 - k_0)}. \qquad (7)$$

Relation (7) shows what the probabilities of earthquakes are falling at the i-th energy interval for the known probability $P(k_0)$, where $\sum_i P(k_i) = 1$ (a complete group of events). In frequency representation, the probability of seismic events falling at the given interval is $P(k_i) = \frac{N(k_i)}{N_\Sigma}$. Therefore, Eq. (7) obtains the following shape:

$$N(k_i) = N(k_0) \cdot 10^{-\gamma(k_i - k_0)} \qquad (8)$$

(N_Σ is the total number of earthquakes recorded in the area S for time T within the energy class interval from k_0 to $k_{\max}$; i.e., $N_\Sigma = \sum_i N(k_i)$. In (8), the following designations were adopted: $N(k_i)$ is the number of earthquakes of the energy class within the interval Δk_i; $N(k_0)$ is the number of earthquakes at the initial interval $k_0 \pm 0.5$. In logarithmic scale, we obtain for (8)

$$\log N(k_i) = \log N(k_0) - \gamma(k_i - k_0). \qquad (9)$$

In expressions (8) and (9), as distinct from (7), there is no normalizing factor $\frac{1}{N_\Sigma}$. Therefore, they are rigidly bound to the number of events $N(k_0)$ within the interval $k_0 \pm 0.5$ in the studied area S for the given T.

Let us express the recurrence pattern in logarithmic form via magnitude M evaluated from surface waves. The M–k relationship for Kamchatka earthquakes is determined from the relation $k = 4.6 + 1.5M$ [5]. Therefore, for (9), we have:

$$\log N(M_i) = \log N(M_0) - 1.5\gamma(M_i - M_0). \qquad (10)$$

Let us introduce the following designations: $A = \log N(M_0) + bM_0 = \log[10^{-bM_0} N(M_0)]$ and $b = 1.5\gamma$, where $M_0 = \frac{k_0 - 4.6}{1.5}$ ($M_0 = 3.27$ at $k_0 = 9.5$, $M_0 = 5.27$ at $k_0 = 12.5$, and $\Delta M = 0.67$ at $\Delta k = 1$. Therefore, the number of events falling at these intervals is the same.) Relation (10) obtains the following shape:

$$\log N = A - bM_i. \qquad (11)$$

Expression (11) is congruent with the Gutenberg–Richter law [4], while the sense of coefficients A and b follows from expressions (9) and (10): A is the logarithm of the number N of seismic events falling at the initial interval $M_0 \pm 0.335$ for the given area S and the time interval T; b is the product of the inclination factor γ and the multiplier 1.5, which characterizes the relationship with magnitude M for the Kamchatka earthquakes of class k [5]. A least-squares computation of b for $k_i \le 12.5$ provides $b_1' = 0.684$ ($\gamma_1 = 0.456$); $b_1'' = 0.679$ ($\gamma_1 = 0.453$) based on Eqs. (5) and (6). In turn, at $k_i \ge 12.5$ $b_2' = 0.925$ (least-squares method, $\gamma_2 = 0.617$) and $b_2'' = 1.018$ (system (5), (6); $\gamma_2 = 0.679$). The values of b_2' and b_2'' are close to unity. Let us calculate $A_{1,2}$ in Eq. (11) for two areas with magnitude $M \le 5.27$ and $M \ge 5.27$. For this purpose, let us determine the number of events falling at the interval $M_0^1 = 3.27 \pm 0.335$: $N_0(M_0^1) = P(3.27 \pm 0.335)N_\Sigma = 0.662 \cdot 21\,845 = 14\,461$ (in our case, for the raw catalogue $N_r = N_\Sigma$ and $f(k_i) = P(k_i)$ at $\Delta k = 1$). And finally we obtain: $A_1 = \log N(M_0) + bM_0 = \log 14461 + 0.684 \cdot 3.27 = 6.397$. Let us calculate A_2. Since $M_0^2 = 5.27 \pm 0.335$, $N_0(M_0^2) = P(5.27 \pm 0.335)N_\Sigma = 0.029 \cdot 21845 = 633$. Thus, $A_2 = \log 633 + 0.9255 \cdot 5.27 = 7.68$.

Let us write down the equations in the Gutenberg–Richter form based on the least-squares computed γ:

$$\log N(M) = 6.397 - 0.684M \quad (M \le 5.27), \qquad (12)$$

$$\log N(M) = 7.68 - 0.925M \quad (M \ge 5.27). \qquad (13)$$

Numerical values of A and b in (12) and (13) correlate well with the respective factors of the Gutenberg–Richter law provided in [4]. However, the derivation of these relations more substantially justifies the explanation of division of the $\log N - M$ relationship into two expressions with distinct definition of the interface conditioned by the recurrence pattern bend, when $M = 5.27$ ($k = 12.5$).

Let us evaluate the maximum magnitude M_{max} of an earthquake intrinsic to the area S_1 for the time period T. As the real number of events must be integer, then adopting it to be equal to 1, we obtain $M_{max} = \frac{7.68}{0.925} = 8.3$. Therefore, we should expect, in general, one earthquake (M 8.3 ± 0.375) in the ~40-yr-long period ($T = 37$ yr) in the area S_1 ($\Delta\varphi = 50°–59°$ N, $\Delta\lambda = 153°–168°$ E).

ACKNOWLEDGMENTS

The author is grateful to the Kamchatka Branch, Geophysical Service, Russian Academy of Sciences (Petropavlovsk-Kamchatski) for the kindly provided earthquake data.

REFERENCES

1. V. N. Gaiskii, *Statistical Studies of Seismic Regime* (Nauka, Moscow, 1970) [in Russian].
2. A. N. Kolmogorov, *Basic Concepts of the Probability Theory* (Nauka, Moscow, 1974) [in Russian].
3. Yu. V. Riznichenko, *Problems of Seismology. Selected Works* (Nauka, Moscow, 1985) [in Russian].
4. Ch. F. Richter, *Elementary Seismology* (W. H. Freeman & Co., San Francisco, 1958; Inostran. Literat., Moscow, 1963).
5. S. A. Fedotov, *Energy Classification of Kuril–Kamchatka Earthquakes and the Problem of Magnitudes* (Nauka, Moscow, 1972) [in Russian].