

## A Bayesian Approach to Seismic Hazard Estimation: Maximum Values of Magnitudes and Peak Ground Accelerations.

V.F.Pisarenko (i), A.A.Lyubushin (ii)

(i): Academy of Sciences of Russia  
 International Institute for Earthquake Prediction  
 Russia, 113556, Moscow, Varshavskoye SH., 79  
 E-mail: [vlad@sirius.mitp.ru](mailto:vlad@sirius.mitp.ru)

(ii): Academy of Sciences of Russia,  
 Institute of the Physics of the Earth,  
 Russia, 123810, Moscow, Bolshaya Gruzinskaya, 10  
 E-mail: [lyubushin@yandex.ru](mailto:lyubushin@yandex.ru)

### Abstract

A rather simple straightforward procedure of estimating maximum values of the considered parameter (earthquake magnitude in a given region or seismic peak ground acceleration at the considered site) and quantiles of its probabilistic distribution on a future time interval of a given length is presented. For peak ground acceleration assessment the input information for the method are seismic catalog and regression relation between peak seismic acceleration at a given point and magnitude and distance from the considered site to epicenter (seismic affect attenuation law). The method is based on Bayesian approach which simply allows to account the influence of uncertainties of magnitudes and seismic acceleration values. The main assumptions for the method are Poissonian character of seismic events flow, a frequency-magnitude law of Gutenberg-Richter's type with cutoff maximal value for estimated parameter and a seismic catalog, having a rather big number of events. The method is applying to seismic hazard estimating in California, Balkans and Japan.

### Methodic.

Let  $\mathbf{R}$  be some value, which was measured or estimated as a sequence on a "past" time interval  $(-\tau, 0)$ :

$$\bar{\mathbf{R}}^{(n)} = (\mathbf{R}_1, \dots, \mathbf{R}_n), \quad \mathbf{R}_i \geq \mathbf{R}_0, \quad \mathbf{R}_\tau = \max_{1 \leq i \leq n} (\mathbf{R}_1, \dots, \mathbf{R}_n) \quad (1)$$

Values (1) could have an arbitrary physical nature. Below we shall consider (1) as earthquakes magnitudes in a given seismoactive region or logarithm of seismic peak ground accelerations at a given site.  $\mathbf{R}_0$  is a minimal cutoff value, i.e. such value which is defined by possibilities of registration systems or was chosen as minimal value up from which values sequence (1) is statistically representative.

*First our assumption* is that values (1) obey the Gutenberg-Richter law of distribution:

$$Prob\{\mathbf{R} < x\} = F(x|\mathbf{R}_0, \rho, \beta) = \frac{e^{-\beta \cdot \mathbf{R}_0} - e^{-\beta \cdot x}}{e^{-\beta \cdot \mathbf{R}_0} - e^{-\beta \cdot \rho}}, \quad \mathbf{R}_0 \leq x \leq \rho \quad (2)$$

Here  $\rho$  is the unknown parameter that has a sense of maximal possible value of  $\mathbf{R}$ , for instance, maximal possible value of earthquake magnitude at a given region. Unknown parameter  $\beta$  usually is called as "slope" of Gutenberg-Richter law at small values of  $x$  when the dependence (2) is plotted in doubly logarithmic axes.

*Second our assumption* is that the sequence (1) is Poissonian process with some intensity value  $\lambda$ , which is unknown parameter also.

Thus the full vector of unknown parameter is the following:

$$\theta = (\rho, \beta, \lambda) \quad (3)$$

For brevity all functions of distribution and statistics of the sequence (1) we shall denote as  $\cdot(\cdot|\boldsymbol{\theta})$ , for example, (2) - as  $\mathbf{F}(\mathbf{x}|\boldsymbol{\theta})$ , argument  $\mathbf{R}_0$  will be omitted.

Probabilistic density of distribution, according to the law (2):

$$\mathbf{f}(\mathbf{x}|\boldsymbol{\theta}) = \mathbf{F}'(\mathbf{x}|\boldsymbol{\theta}) = \frac{\beta \cdot e^{-\beta \cdot \mathbf{x}}}{e^{-\beta \cdot \mathbf{R}_0} - e^{-\beta \cdot \rho}} \quad (4)$$

Let's introduce now an error  $\boldsymbol{\varepsilon}$ , with which we know values (1), i.e. for us really in (1) are accessible not *true*, but *apparent* values of  $\mathbf{R}$ , which are defined by formula:

$$\tilde{\mathbf{R}} = \mathbf{R} + \boldsymbol{\varepsilon} \quad (5)$$

and let  $\mathbf{n}(\mathbf{x}|\boldsymbol{\delta})$  be a density of probabilistic distribution of the error  $\boldsymbol{\varepsilon}$ , where  $\boldsymbol{\delta}$  is a given scale parameter of the density. We shall use below a uniform distribution density:

$$\begin{aligned} \mathbf{n}(\mathbf{x}|\boldsymbol{\delta}) &= \frac{1}{2\boldsymbol{\delta}}, \quad |\mathbf{x}| \leq \boldsymbol{\delta} \\ &= 0, \quad |\mathbf{x}| > \boldsymbol{\delta} \end{aligned} \quad (6)$$

Then a distribution density of the apparent values is the following:

$$\tilde{\mathbf{f}}(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\delta}) = \int_{-\infty}^{+\infty} \mathbf{f}(\xi|\boldsymbol{\theta}) \mathbf{n}(\mathbf{x} - \xi|\boldsymbol{\delta}) d\xi = \frac{\mathbf{F}(\mathbf{x} + \boldsymbol{\delta}|\boldsymbol{\theta}) - \mathbf{F}(\mathbf{x} - \boldsymbol{\delta}|\boldsymbol{\theta})}{2\boldsymbol{\delta}} \quad (7)$$

Let  $\tilde{\mathbf{F}}(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\delta})$  be a function of distribution which is correspondent to the density (7). Because  $\mathbf{F}(\mathbf{x}|\boldsymbol{\theta}) = 0$  for  $\mathbf{x} < \mathbf{R}_0$ , then

$$\tilde{\mathbf{F}}(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\delta}) = \int_{\mathbf{R}_0 - \boldsymbol{\delta}}^{\mathbf{x}} \tilde{\mathbf{f}}(\xi|\boldsymbol{\theta}, \boldsymbol{\delta}) d\xi \quad (8)$$

As apparent values  $\tilde{\mathbf{R}} \geq \mathbf{R}_0$  and  $\tilde{\mathbf{f}}(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\delta}) > 0$  for  $\mathbf{x} \in (\mathbf{R}_0 - \boldsymbol{\delta}, \mathbf{R}_0)$  then we shall renormalize  $\tilde{\mathbf{f}}(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\delta})$  and  $\tilde{\mathbf{F}}(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\delta})$  in such a way that they will equal zero for  $\mathbf{x} < \mathbf{R}_0$ :

$$\begin{aligned} \bar{\mathbf{f}}(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\delta}) &= \frac{\tilde{\mathbf{f}}(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\delta})}{1 - \tilde{\mathbf{F}}(\mathbf{R}_0|\boldsymbol{\theta}, \boldsymbol{\delta})}, \quad \text{for } \mathbf{x} \geq \mathbf{R}_0 \\ &= 0, \quad \text{for } \mathbf{x} < \mathbf{R}_0 \end{aligned} \quad (9)$$

Function of distribution, which corresponds to the density (9), is defined by formula:

$$\begin{aligned} \bar{\mathbf{F}}(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\delta}) &= \frac{\tilde{\mathbf{F}}(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\delta}) - \tilde{\mathbf{F}}(\mathbf{R}_0|\boldsymbol{\theta}, \boldsymbol{\delta})}{1 - \tilde{\mathbf{F}}(\mathbf{R}_0|\boldsymbol{\theta}, \boldsymbol{\delta})}, \quad \text{for } \mathbf{x} \geq \mathbf{R}_0 \\ &= 0, \quad \text{for } \mathbf{x} < \mathbf{R}_0 \end{aligned} \quad (10)$$

Now we want to derive a relationship between intensity  $\boldsymbol{\lambda}$  of "true"  $\mathbf{R}$ -values and intensity  $\bar{\boldsymbol{\lambda}}$  of their apparent values. As  $\mathbf{R} = \tilde{\mathbf{R}} - \boldsymbol{\varepsilon}$  and  $-\boldsymbol{\varepsilon}$  is also distributed according to (6), then

$$\mathbf{f}(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\delta}) = \int_{-\infty}^{+\infty} \bar{\mathbf{f}}(\xi|\boldsymbol{\theta}, \boldsymbol{\delta}) \cdot \mathbf{n}(\mathbf{x} - \xi|\boldsymbol{\delta}) d\xi \quad (11)$$

A share of those apparent  $\mathbf{R}$  -values, for which true values  $< \mathbf{R}_0$  equals:

$$\kappa = \int_{-\infty}^{\mathbf{R}_0} \mathbf{f}(\mathbf{x}|\boldsymbol{\theta}) \mathbf{d}\mathbf{x} = \int_{-\infty}^{\mathbf{R}_0} \int_{-\infty}^{+\infty} \bar{\mathbf{f}}(\xi|\boldsymbol{\theta}, \delta) \cdot \mathbf{n}(\mathbf{x} - \xi|\delta) \mathbf{d}\xi \mathbf{d}\mathbf{x} \quad (12)$$

Then, in the assumption of Poissonian character of the sequence (1) it follows that:

$$\lambda = \bar{\lambda} \cdot (1 - \kappa) \quad (13)$$

Substituting (7) into (12) and using the fact that  $\bar{\mathbf{F}}=0$  for  $\mathbf{x}<\mathbf{R}_0$  we'll obtain:

$$\kappa = \int_{-\infty}^{\mathbf{R}_0} \frac{\bar{\mathbf{F}}(\mathbf{x} + \delta|\boldsymbol{\theta}, \delta) - \bar{\mathbf{F}}(\mathbf{x} - \delta|\boldsymbol{\theta}, \delta)}{2\delta} \mathbf{d}\mathbf{x} = \frac{1}{2\delta} \int_{\mathbf{R}_0 - \delta}^{\mathbf{R}_0} \bar{\mathbf{F}}(\mathbf{x} + \delta|\boldsymbol{\theta}, \delta) \mathbf{d}\mathbf{x} = \frac{1}{2\delta} \int_{\mathbf{R}_0}^{\mathbf{R}_0 + \delta} \bar{\mathbf{F}}(\mathbf{x}|\boldsymbol{\theta}, \delta) \mathbf{d}\mathbf{x}$$

Thus:

$$\bar{\lambda} = \bar{\lambda}(\boldsymbol{\theta}, \delta) = \frac{\lambda}{1 - \frac{1}{2\delta} \int_{\mathbf{R}_0}^{\mathbf{R}_0 + \delta} \bar{\mathbf{F}}(\mathbf{x}|\boldsymbol{\theta}, \delta) \mathbf{d}\mathbf{x}} \quad (14)$$

As it was shown in [Bender, 1987], see also [Kijko, Sellevoll, 1989, 1992], formulae (9), (10) and (14) could be simplified. Let's introduce values:

$$\mathbf{A}(\mathbf{x}) = \exp(-\boldsymbol{\beta}\mathbf{x}), \mathbf{A}_1 = \mathbf{A}(\mathbf{R}_0), \mathbf{A}_2 = \mathbf{A}(\boldsymbol{\rho}), \mathbf{c}_f = \mathbf{c}_f(\boldsymbol{\theta}, \delta) = \frac{\exp(\boldsymbol{\beta}\delta) - \exp(-\boldsymbol{\beta}\delta)}{2\delta}$$

Then:

$$\begin{aligned} \bar{\mathbf{f}}(\mathbf{x}|\boldsymbol{\theta}, \delta) &= \frac{1}{(\mathbf{c}_f \mathbf{A}_1 - \mathbf{A}_2)} \cdot \hat{\mathbf{f}}, \hat{\mathbf{f}} = \mathbf{c}_f \boldsymbol{\beta} \mathbf{A}(\mathbf{x}), \text{ for } \mathbf{R}_0 \leq \mathbf{x} < \boldsymbol{\rho} - \delta \\ \hat{\mathbf{f}} &= \frac{\mathbf{A}(\mathbf{x} - \delta) - \mathbf{A}_2}{2\delta}, \text{ for } \boldsymbol{\rho} - \delta \leq \mathbf{x} \leq \boldsymbol{\rho} + \delta \end{aligned} \quad (15)$$

and:

$$\bar{\mathbf{F}}(\mathbf{x}|\boldsymbol{\theta}, \delta) = \frac{1}{(\mathbf{c}_f \mathbf{A}_1 - \mathbf{A}_2)} \cdot \hat{\mathbf{F}}, \hat{\mathbf{F}} = \mathbf{c}_f (\mathbf{A}_1 - \mathbf{A}(\mathbf{x})), \text{ for } \mathbf{R}_0 \leq \mathbf{x} < \boldsymbol{\rho} - \delta \quad (16)$$

$$\hat{\mathbf{F}} = \mathbf{c}_f (\mathbf{A}_1 - \mathbf{A}(\boldsymbol{\rho} - \delta)) - \mathbf{A}_2 \cdot \frac{(\mathbf{x} - \boldsymbol{\rho} + \delta)}{2\delta} - \frac{[\mathbf{A}(\mathbf{x} - \delta) - \mathbf{A}(\boldsymbol{\rho} - 2\delta)]}{2\boldsymbol{\beta}\delta}, \text{ for } \boldsymbol{\rho} - \delta \leq \mathbf{x} \leq \boldsymbol{\rho} + \delta$$

$$\bar{\lambda}(\boldsymbol{\theta}, \delta) = \hat{\lambda} \cdot \mathbf{c}_f(\boldsymbol{\theta}, \delta) \quad (17)$$

Now we are ready to proceed directly to Bayesian estimates. Let  $\boldsymbol{\Pi}$  be *a priori* uncertainty domain of values of parameters  $\boldsymbol{\theta}$ :

$$\boldsymbol{\Pi} = \{ \lambda_{\min} \leq \lambda \leq \lambda_{\max}, \boldsymbol{\beta}_{\min} \leq \boldsymbol{\beta} \leq \boldsymbol{\beta}_{\max}, \boldsymbol{\rho}_{\min} \leq \boldsymbol{\rho} \leq \boldsymbol{\rho}_{\max} \} \quad (18)$$

We shall consider *a priori* density of the vector  $\boldsymbol{\theta}$  to be uniform in the domain  $\boldsymbol{\Pi}$ .

Let  $[0, T]$  be a future interval of time for which we want to estimate function of distribution of maximal value  $\rho$  and its quantiles.

As the flow of events (1) is stationary and Poissonian then it follows that intensity of event with  $R < x$  equals  $\lambda \cdot F(x|\theta)$  and intensity of events with  $R \geq x$  equals  $\lambda \cdot (1 - F(x|\theta))$ . From Poissonian character of the sequence (1) it follows that probability that it will be no events with  $R \geq x$  on time interval  $[0, T]$  or that all events on  $[0, T]$  will have  $R < x$  equals:

$$\exp(-\lambda \cdot (1 - F(x|\theta)) \cdot T) \quad (19)$$

Let's denote by  $R_T$  maximal value of  $R$  on the time interval  $[0, T]$ . Then  $Prob\{R_T < x\} = \exp(-\lambda \cdot (1 - F(x|\theta)) \cdot T)$ . But into this probability a case when there are no events on  $[0, T]$  is included also. Let's denote by  $v_T$  the number of events with  $R \geq R_0$  on the interval  $[0, T]$ . Then

$$Prob\{v_T = 0\} = e^{-\lambda \cdot T}; \quad Prob\{v_T \geq 1\} = (1 - e^{-\lambda \cdot T})$$

That is why:

$$\begin{aligned} \Phi_T(x|\theta) &= Prob\{R_T < x | v_T \geq 1\} = \\ &= \frac{\exp(-\lambda T(1 - F(x|\theta))) - \exp(-\lambda T)}{1 - \exp(-\lambda T)} = \frac{\exp(\lambda T F(x|\theta)) - 1}{\exp(\lambda T) - 1} \end{aligned} \quad (20)$$

Formula (20) defines an expression for *a priori* function of distribution for *true* maximal values of  $R$  on time interval  $[0, T]$ . Let's introduce also the following functions:

$$\phi_T(x|\theta) = \frac{d}{dx} \Phi_T(x|\theta) \quad (21)$$

- *a priori* density for *true* maximal values of  $R$  on time interval  $[0, T]$ ;

$$Y_T(\alpha|\theta) - \text{a root of equation: } \Phi_T(x|\theta) = \alpha, \quad 0 \leq \alpha \leq 1 \quad (22)$$

- *a priori* quantile for probability  $\alpha$  for *true* maximal values of  $R$  on time interval  $[0, T]$ ;

If we substitute in formula (20)  $F(x|\theta) \rightarrow \bar{F}(x|\theta, \delta)$  then we'll obtain a function:

$\bar{\Phi}_T(x|\theta, \delta)$  - *a priori* function of distribution for *apparent* maximal values of  $R$  on time interval  $[0, T]$ .

Substituting  $\bar{\Phi}_T(x|\theta, \delta)$  into formulae (21) and (22), we'll obtain:

$\bar{\phi}_T(x|\theta, \delta)$  - *a priori* density for *apparent* maximal values of  $R$  on time interval  $[0, T]$  and:

$\bar{Y}_T(\alpha|\theta, \delta)$  - *a priori* quantile for probability  $\alpha$  for *apparent* maximal values of  $R$  on time interval  $[0, T]$ .

According to definition of conditional probability, *a posterior* density of distribution of vector of parameters  $\theta$  equals to:

$$f(\theta|\bar{R}^{(n)}, \delta) = \frac{f(\theta, \bar{R}^{(n)}|\delta)}{f(\bar{R}^{(n)}|\delta)} \quad (23)$$

but  $f(\theta, \bar{R}^{(n)}|\delta) = f(\bar{R}^{(n)}|\theta, \delta) \cdot f^a(\theta)$ , where  $f^a(\theta)$  is *a priori* density of distribution of vector  $\theta$  in the domain  $\Pi$ . As  $f^a(\theta) = \text{const}$  according to our assumption and taking into consideration that:

$$f(\bar{R}^{(n)}|\delta) = \int_{\Pi} f(\bar{R}^{(n)}|\theta, \delta) d\theta$$

we'll obtain after using a Bayes formula [Rao, 1965] and normalizing the density that:

$$\mathbf{f}(\boldsymbol{\theta}|\bar{\mathbf{R}}^{(n)}, \boldsymbol{\delta}) = \frac{\mathbf{f}(\bar{\mathbf{R}}^{(n)}|\boldsymbol{\theta}, \boldsymbol{\delta})}{\int_{\Pi} \mathbf{f}(\bar{\mathbf{R}}^{(n)}|\boldsymbol{\vartheta}, \boldsymbol{\delta}) d\boldsymbol{\vartheta}} \quad (24)$$

Formula (24) is *our main formula for computing a posterior density of distribution of vector of parameters*  $\boldsymbol{\theta}$ . In order to use (24) we must have an expression for the function  $\mathbf{f}(\bar{\mathbf{R}}^{(n)}|\boldsymbol{\theta}, \boldsymbol{\delta})$ . Having the assumption of Poissonian character of the sequence (1) and of independency of its members, we can obtain:

$$\mathbf{f}(\bar{\mathbf{R}}^{(n)}|\boldsymbol{\theta}, \boldsymbol{\delta}) = \bar{\mathbf{f}}(\mathbf{R}_1|\boldsymbol{\theta}, \boldsymbol{\delta}) \cdots \bar{\mathbf{f}}(\mathbf{R}_n|\boldsymbol{\theta}, \boldsymbol{\delta}) \cdot \frac{\exp(-\bar{\lambda}(\boldsymbol{\theta}, \boldsymbol{\delta}) \cdot \tau) \cdot (\bar{\lambda}(\boldsymbol{\theta}, \boldsymbol{\delta}) \cdot \tau)^n}{n!} \quad (25)$$

Now we are ready completely to compute a Bayesian estimate of vector  $\boldsymbol{\theta}$ :

$$\hat{\boldsymbol{\theta}}(\bar{\mathbf{R}}^{(n)}|\boldsymbol{\delta}) = \int_{\Pi} \boldsymbol{\vartheta} \cdot \mathbf{f}(\boldsymbol{\vartheta}|\bar{\mathbf{R}}^{(n)}, \boldsymbol{\delta}) d\boldsymbol{\vartheta} \quad (26)$$

Among one of its component vector (26) contains an estimate of maximal value  $\boldsymbol{\rho}$ . Using analogous to (26) formulae, we can obtain Bayesian estimates of any of the functions (20), (21), (22). The most interesting for us are estimates of quantiles of functions of distribution of true and apparent  $\mathbf{R}$ -values on a given future time interval  $[\mathbf{0}, \mathbf{T}]$ , for instance for  $\boldsymbol{\alpha}$ -quantiles of apparent values:

$$\hat{\bar{\mathbf{Y}}}_T(\boldsymbol{\alpha}|\bar{\mathbf{R}}^{(n)}, \boldsymbol{\delta}) = \int_{\Pi} \bar{\mathbf{Y}}_T(\boldsymbol{\alpha}|\boldsymbol{\vartheta}, \boldsymbol{\delta}) \cdot \mathbf{f}(\boldsymbol{\vartheta}|\bar{\mathbf{R}}^{(n)}, \boldsymbol{\delta}) d\boldsymbol{\vartheta} \quad (27)$$

$\hat{\bar{\mathbf{Y}}}_T(\boldsymbol{\delta}|\bar{\mathbf{R}}^{(n)}, \boldsymbol{\delta})$  for  $\boldsymbol{\alpha}$ -quantiles of true values is written analogously to (27). Using averaging over the density (24), (25) we can estimate also variances of Bayesian estimates (26), (27). For example

$$\text{var}\{ \hat{\bar{\mathbf{Y}}}_T(\boldsymbol{\alpha}|\bar{\mathbf{R}}^{(n)}, \boldsymbol{\delta}) \} = \int_{\Pi} (\bar{\mathbf{Y}}_T(\boldsymbol{\alpha}|\boldsymbol{\vartheta}, \boldsymbol{\delta}) - \hat{\bar{\mathbf{Y}}}_T(\boldsymbol{\alpha}|\bar{\mathbf{R}}^{(n)}, \boldsymbol{\delta}))^2 \cdot \mathbf{f}(\boldsymbol{\vartheta}|\bar{\mathbf{R}}^{(n)}, \boldsymbol{\delta}) d\boldsymbol{\vartheta} \quad (28)$$

In order to finish description of the methodic, we must define the domain of *a priori* uncertainty  $\Pi$  (18).

First of all we set  $\boldsymbol{\rho}_{\min} = \mathbf{R}_\tau - \boldsymbol{\delta}$ . As for value of  $\boldsymbol{\rho}_{\max}$ , it is introduced by the user of the method and depends of the specifics of the data series (1). For instance for estimation of maximal magnitudes in California we put  $\boldsymbol{\rho}_{\max} = 9$ . Boundary values for the slope  $\boldsymbol{\beta}$  are defined by formulae:

$$\boldsymbol{\beta}_{\min} = \boldsymbol{\beta}_0 \cdot (1 - \gamma), \quad \boldsymbol{\beta}_{\max} = \boldsymbol{\beta}_0 \cdot (1 + \gamma), \quad 0 < \gamma \leq 1 \quad (29)$$

where  $\boldsymbol{\beta}_0$  is the "central" value, obtained as a maximum likelihood estimate of the slope for Gutenberg-Richter law:

$$\sum_{i=1}^n \ln \left\{ \frac{\boldsymbol{\beta} \cdot e^{-\boldsymbol{\beta} \cdot \mathbf{R}_i}}{e^{-\boldsymbol{\beta} \cdot \mathbf{R}_0} - e^{-\boldsymbol{\beta} \cdot \mathbf{R}_\tau}} \right\} \rightarrow \max_{\boldsymbol{\beta}, \boldsymbol{\beta} \in (0, \boldsymbol{\beta}_s)} \quad (30)$$

Here  $\boldsymbol{\beta}_s$  is a rather big value, for example 10, value  $\gamma$  is a parameter of the method, usually we take  $\gamma = 0.5$ .

For setting boundary values for intensity in (18) we use the following reasons. As a consequence of normal approximation for Poissonian process for rather big  $n$  [Cox, Lewis, 1966], variance of the value  $\boldsymbol{\lambda}\tau$  has approximate value  $\sqrt{n} \approx \sqrt{\boldsymbol{\lambda}\tau}$ . So taking boundaries  $\pm 3\sigma$ , we'll obtain:

$$\lambda_{\min} = \lambda_0 \cdot \left(1 - \frac{3}{\sqrt{\lambda_0 \tau}}\right), \quad \lambda_{\max} = \lambda_0 \cdot \left(1 + \frac{3}{\sqrt{\lambda_0 \tau}}\right) \quad (31)$$

$$\text{where } \lambda_0 = \frac{\bar{\lambda}_0}{c_f(\beta_0, \delta)}, \quad \bar{\lambda}_0 = \frac{n}{\tau}.$$

Thus the description of methodic is finished completely.

### Examples of applications.

In papers [Pisarenko, Lyubushin et al., 1996; Pisarenko, Lyubushin, 1997] we apply the presented above methodic to estimating maximal values of magnitudes and seismic peak ground accelerations, their functions of distribution and quantiles for a number of regions in California, Italy and Caucasus.

Now we shall emphasize on the using of the method to estimation of maximal values of seismic peak ground accelerations at a given site for other regions. Note that problem for maximal peak ground accelerations essentially differs from those concerning maximal magnitudes. First of all, direct measurements of seismic accelerations are very rare and fragmental. That is why there is no catalogs, containing values of maximal accelerations for most interesting sites, but there are a lot of so called “attenuation laws”, which represent some functions between logarithm of maximal accelerations  $\mathbf{R} = \lg(\mathbf{A}_{\max})$  and magnitude  $\mathbf{M}$  of the earthquake and distance  $\mathbf{r}$  from the considered site to epicenter of the earthquake:

$$\mathbf{R} = \lg(\mathbf{A}_{\max}) = \Psi(\mathbf{M}, \mathbf{r}) \quad (32)$$

Usually functions (32) are empirical regression laws, obtained by collecting data from a specified region and fitting to them some class of functions. In our calculations we use the following forms of (32):

$$\begin{aligned} \lg(\mathbf{A}_{\max}) &= 0.28\mathbf{M} - 0.8\lg(\mathbf{r}) + 1.7 \quad \text{for } \mathbf{A}_{\max} > 160 \text{ cm/sec}^2 \\ &= 0.80\mathbf{M} - 2.3\lg(\mathbf{r}) + 0.8 \quad \text{for } \mathbf{A}_{\max} \leq 160 \text{ cm/sec}^2 \end{aligned} \quad (32a)$$

- Aptikaev's law, the whole world [Steinberg et al., 1993];

$$\lg(\mathbf{A}_{\max}) = 0.49 + 0.23(\mathbf{M} - 6) - \lg(\mathbf{D}) - 0.0027\mathbf{D} + \lg(981) \quad (32b)$$

where  $\mathbf{D}^2 = \mathbf{r}^2 + 64$ ; - Joyner-Boor's law, California, hard rock [Joyner, Boore, 1981];

$$\lg(\mathbf{A}_{\max}) = 0.41\mathbf{M} - \lg(\mathbf{r} + 0.032 \cdot 10^{0.41\mathbf{r}}) - 0.0034\mathbf{r} + 1.30 \quad (32c)$$

- Fukushima-Tanoak's law, Japan, hard rock [Fukushima, Tanaka, 1990]. In all of the formulae (32a,b,c)  $\mathbf{r}$  is measured in kilometers,  $\mathbf{A}_{\max}$  - in  $\text{cm/sec}^2$ .

From formula (32a) it can be seen that if  $\mathbf{r}$  is close to zero an outlier peculiarity arises. This defect of the law (32a) was improved by statistical regularization which is connected to the fact that epicenter of the earthquake could not be defined exactly. If  $\mathbf{x}$ ,  $\mathbf{y}$  and  $\mathbf{z}$  coordinates of the epicenter are defined with independent Gaussian errors, having a zero mean and the same standard deviation  $\sigma$ , then

$$\mathbf{E}\{\lg(\mathbf{r}^2)\} = \lg(\sigma^2) + \mathbf{E}\{\lg(\mathbf{q})\}$$

where  $\mathbf{q}$  is a random value with noncentral  $\chi^2$ -distribution having 3 degrees of freedom and noncentral parameter  $\mathbf{r}^2/\sigma^2$  [Kendall, Stuart, 1961]. It can be approximated by a simple relation:

$$\mathbf{E}\{\lg(\mathbf{r}^2)\} \approx \lg(\sigma^2) + 0.7, \quad \text{for } \mathbf{r}/\sigma < 1.4$$

$$\approx \lg(\mathbf{r}^2), \text{ for } \mathbf{r}/\sigma \geq 1.4 \quad (33)$$

When using the formula (32a) we substitute  $2\lg(\mathbf{r})$  on (33) for  $\sigma = 10$  km.

Thus, the sequence (1) is composed of values, computed in accordance with one of the formulae (32). As regression formulae (32) give their “own” errors due to errors of statistical fitting, the “general” error  $\epsilon$  is composed of two part: “own” error and error due to no adequateness of choused class of functions in (32). We suppose that this general error has a zero mean and is distributed uniformly. We must also keep in mind that a “real” relation of the type (32) must be not stationary and be dependent of soil and rock conditions, for example of precipitations intensity to the moment of earthquake. That is why we choose a value of  $\delta$  for uncertainty of  $\lg(\mathbf{A}_{\max})$  is equal to **0.75** in our calculations. This value is approximately twice more than given standard deviations of fitting for formulae (32a,b,c).

Then for satisfying to the assumption of Poissonian character of the events flow (1) we must remove aftershocks from the processed seismic catalogs. This operation for the problem of maximal seismic accelerations has some specific if compare it to usual aftershocks removing. We could not leave only “usual” mainshock because an event-aftershock with less value of magnitude, which has an epicenter more close to the considered site, could generates a more large peak acceleration than “usual” mainshock. That is why an aftershocks removing procedure [Gardner, Knopoff, 1974] was modified in a following way: all events were usually divided into mainshocks and aftershocks and than among each mainshock-aftershocks sequence only one event was leaved - those which generates the most large value of peak acceleration in accordance with used formula (32).

On Figures 1-4 there are results of applying the methods to the following sites:

Fig.1 - San-Francisco, North California, formula (32c);

Fig.2 - Los-Angeles, South California, formula (32c);

Fig.3 - Kobe, Japan, formula (32b);

Fig.4 - Thessaloniki, Greece, formula (32a).

Seismic catalogs were taken from CD-ROMs: [Global hypocenters data base, 1989, 1989-1990, 1996]. In all cases a “past” time interval  $(-\tau, 0)$  which corresponds to period 1900-1995 was used; quantiles of true values of  $(\lg(\mathbf{A}_{\max}))_{\max}$  for  $\alpha = \{0.5, 0.6, 0.7, 0.8, 0.9, 0.95\}$  were estimated on a number of future time intervals  $[0, T]$ , where  $T = \{5, 10, 20, 30, 50 \text{ years}\}$ . Value of  $\rho_{\max}$  was set to be **3.5**.

### Conclusion.

A proposed method for estimating probabilistic characteristics of maximal values of seismic peak ground acceleration on a given future time interval for a given site is rather simple in use and needs only those information, which is in seismic catalogs and used attenuation laws. We do not need previously estimate maximal magnitudes. This method could be useful for “express” estimating in elaborating a detailed maps of seismic hazard using not only seismic information but geological and paleoseismical as well.

### References.

- Bender B. (1987) Effects of observational errors in relating magnitude scales and fitting the Gutenberg-Richter parameter  $\beta$ . Bull. Seism. Soc. Amer. V.77. 1400-1428.
- Cox D.R., Lewis P.A.W. (1966) The statistical analysis of series of events. London, Methuen.
- Fukushima Y., T.Tanaka. (1990) A new relation for peak horizontal acceleration of strong earthquake ground motion in Japan. Bull. Seism. Soc. America. V.80. No.4. 757-783.
- Gardner J.K., L.Knopoff. (1974) Is the sequence of earthquakes in Southern California with aftershocks removed, Poissonian? Bull. Seism. Soc. Amer. V.64. 1363-1367.
- Global hypocenters data base, CD-ROM, 1989, NEIC-USGS; 1989-1990, HDS NEIC, USGS; NOAA, 1996, Golden Co., USA.
- Joyner W., D.Boore. (1981) Peak horizontal acceleration and velocity from strong-motion records including records from the 1979 Imperial Valley, California, earthquake. Bull. Seism. Soc. America. V.71, 2011-2038.
- Kendall M., A.Stuart. (1961) The advanced theory of statistics. Vol.2., Inference and Relationships. London, Griffin.
- Kijko A., M.A.Sellevoll. Estimation of earthquake hazard parameters from incomplete data files.

- Part I.* (1989) Utilization of extreme and complete catalogs with different threshold magnitudes. Bull. Seism. Soc. Amer. V.79. 645-654.
- Part II.* (1992) Incorporation of magnitude heterogeneity. Bull. Seism. Soc. Amer. V.82. 120-134.
- Pisarenko V.F., A.A.Lyubushin, V.B.Lysenko, T.V.Golubeva.* (1996) Statistical estimation of seismic hazard parameters: maximum possible magnitude and related parameters. Bull. Seism. Soc. America. V.86. No.3. 691-700.
- Pisarenko V.F., A.A.Lyubushin.* (1997) Statistical estimation of maximal peak ground acceleration at a given point of seismic region. - Journal of Seismology. Vol.1, pp.395-405.
- Rao C.R.* (1965) Linear statistical inference and its application. New York. John Wiley.
- Steinberg V.V., M.V.Saks, F.F.Aptikaev et al.* (1993) Methods of seismic ground motion estimation (Handbook). In Voprosy Inzhenernoi Seismologii. Iss.34. Moscow, Nauka. 5-97 (in Russian).



**Fig.1. San-Francisco.**

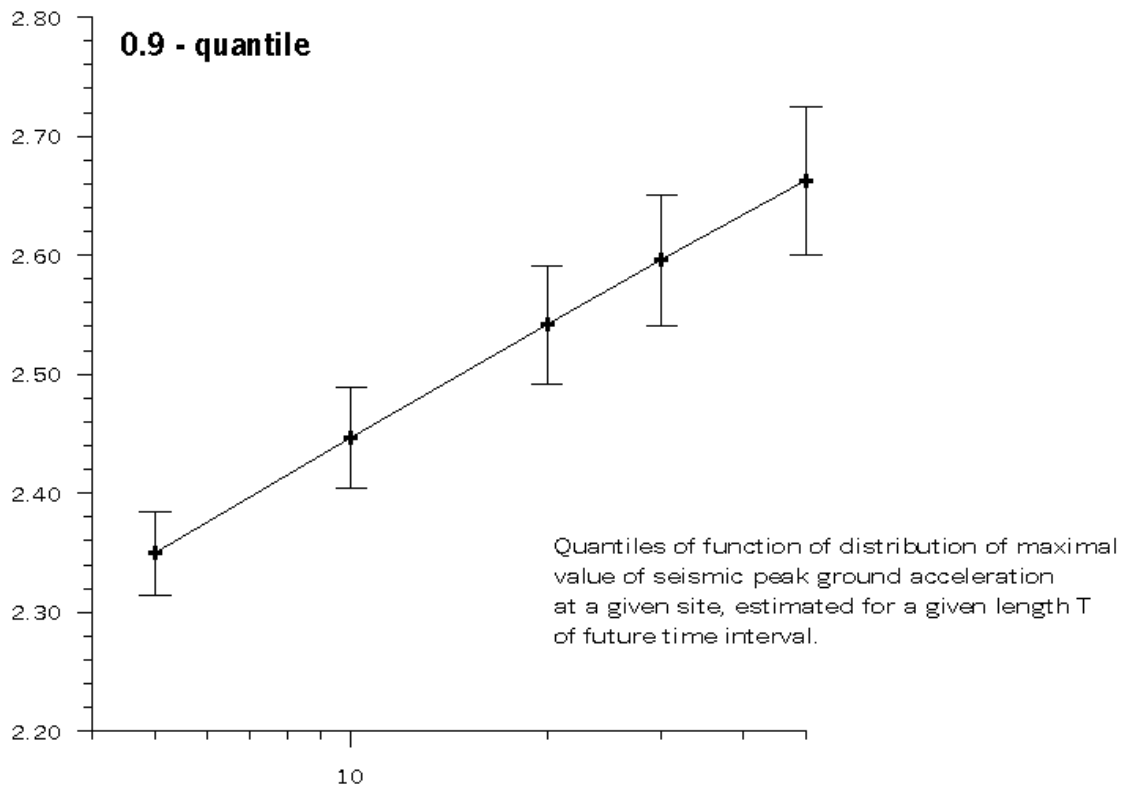
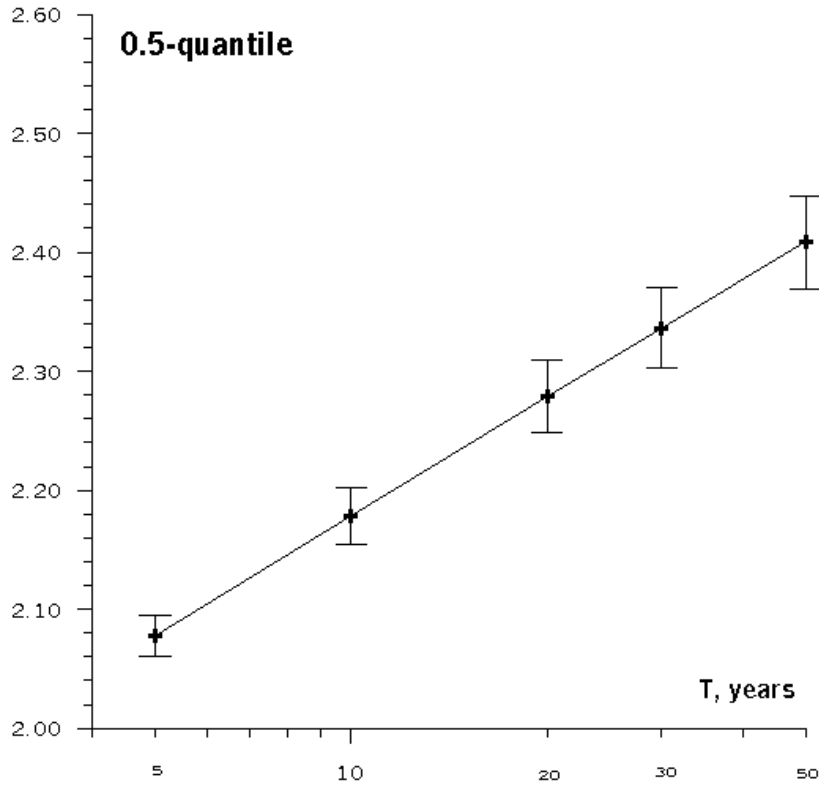


Fig.2. Los-Angeles.

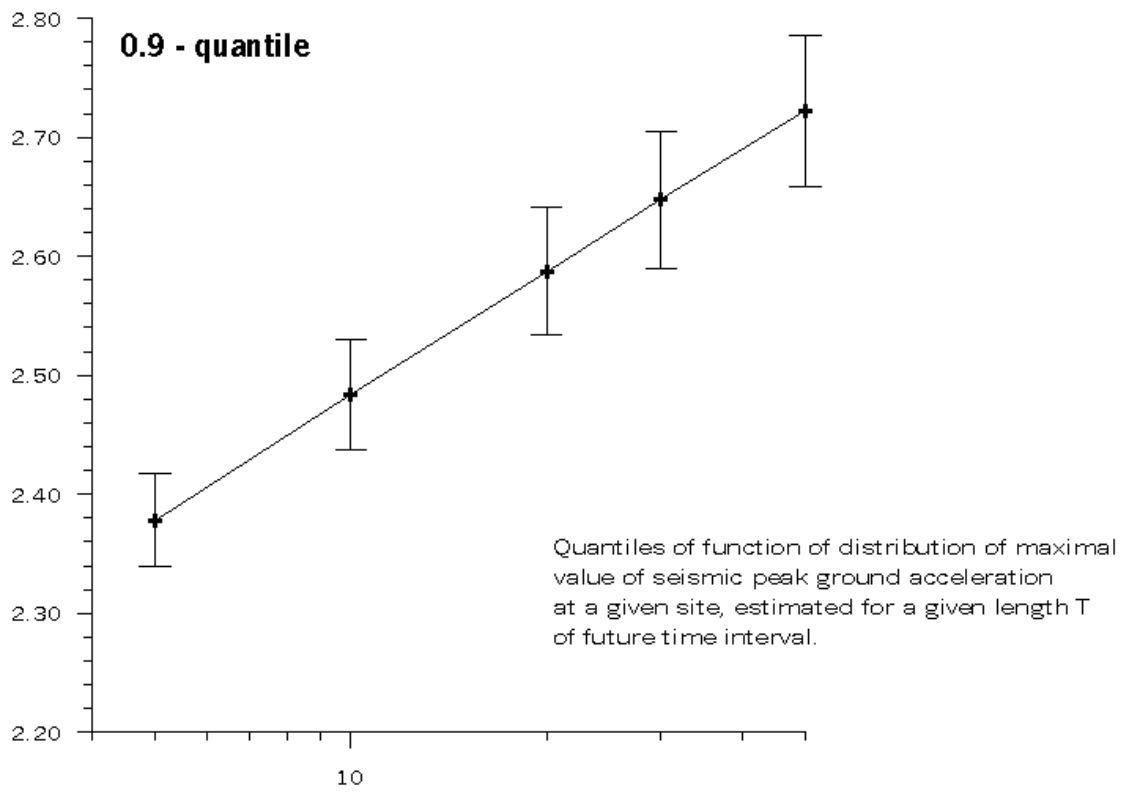
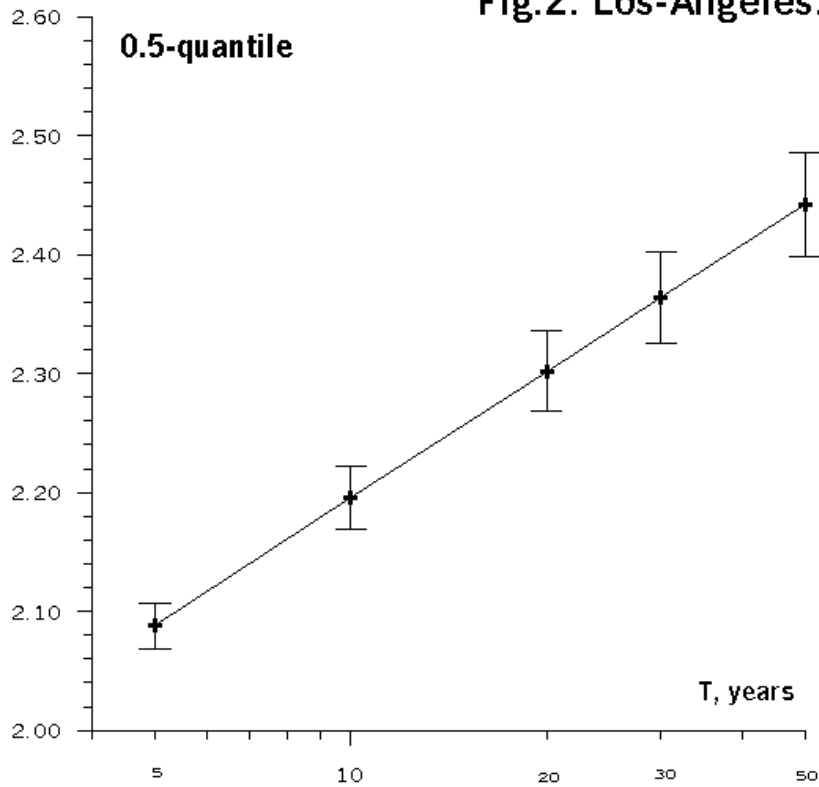
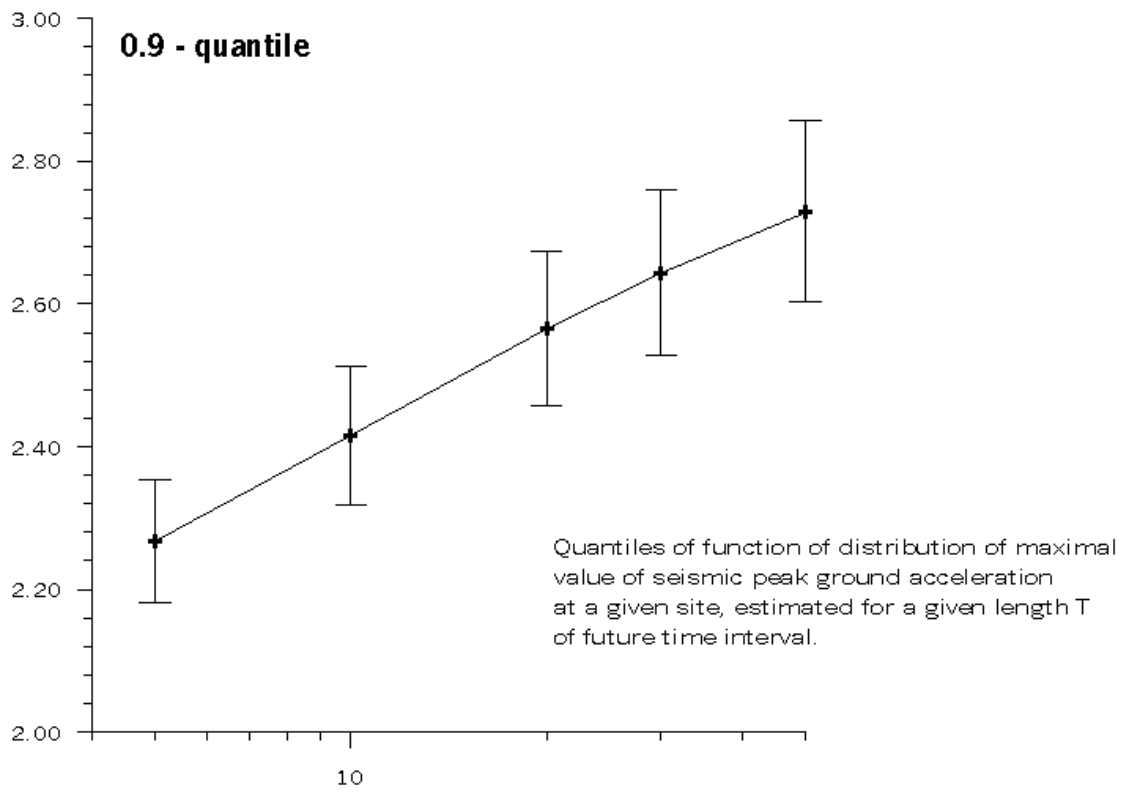
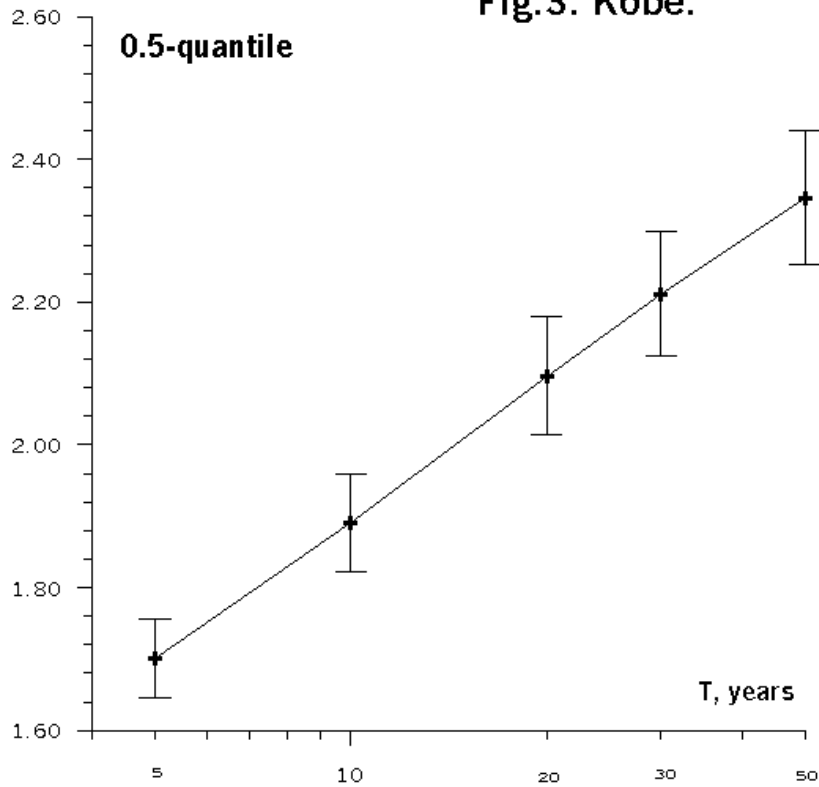


Fig.3. Kobe.



**Fig.4. Thessaloniki.**

