Modeling of nonstationary time series using nonparametric methods

Fedorov Sergei Leonidovich

Basic concepts

assumption 1. Time series values **x**(t) are uniformly bounded in time and belong to the interval [0;1]

defenition 1. SDF F(**x**,t; N) – selective distribution function of the time series fragment $\{x(t-N+1),\ldots,x(t)\}$

defenition 2. $\rho(t;N)=\sup\left|F(x,t;N)-F(x,t+N;N)\right|$ - distance between two samples of length N as norm C^{x}

defeniton 3. G(ρ,N) – distribution function of distances between two samples of length N

defenition 4. SDFD f(**x**,t; N) – selective distribution function density of the time series fragment $\{x(t-N+1),\ldots,x(t)\}$

SDFD as a Histogram

Let SDFD is a histogram uniformly divided into n class intervals, within which the distribution is assumed to be uniform. Then

$$f(x) = f_j, \quad x \in \left[\frac{j}{n}; \frac{j+1}{n}\right], \quad j = 0 \div n-1$$

$$F(x) = (nx - j)f_{j+1} + \sum_{k=1}^{j} f_k, \quad x \in \left[\frac{j}{n}; \frac{j+1}{n}\right], \quad j = 0 \div n-1$$

Solved problems

- Developing of a nonparametric indicator of a breakdown for a selective distribution function in a sliding window;
- Creating of a model of distribution function evolution using the empirical kinetic equation;
- Developing of the method of stochastic process trajectories set generation.

Practical use

- Earthquake research
- Medicine
- Text analysis
- Telecommunications
- Stocks market

Why is it important to take in account the non stationary nature of the series

All theorems for estimating the confidence interval are proved only for the stationary case

The classical theorems on convergence

T1. (Glivenko) Selective disribution $F_N(x)$ of a random stationary quantity uniformly with respect to x converges to the distribution of the general population $F(x)$:

$$P\left\{\lim_{N\to\infty} \sup_x \left|F_N(x) - F(x)\right| = 0\right\} = 1$$

T2. (Kolmogorov) If the general distribution $F(x)$ is continuous, than the statistic

$$\sqrt{N} \sup_x \left|F_N(x) - F(x)\right|$$

converges to the Kolmogorov function:

$$\lim_{N\to\infty} P\left\{0 < \sqrt{N} \sup_x \left|F_N(x) - F(x)\right| < z\right\} = K(z) = \sum_{k=-\infty}^{\infty} (-1)^k \exp\left(-2k^2 z^2\right)$$

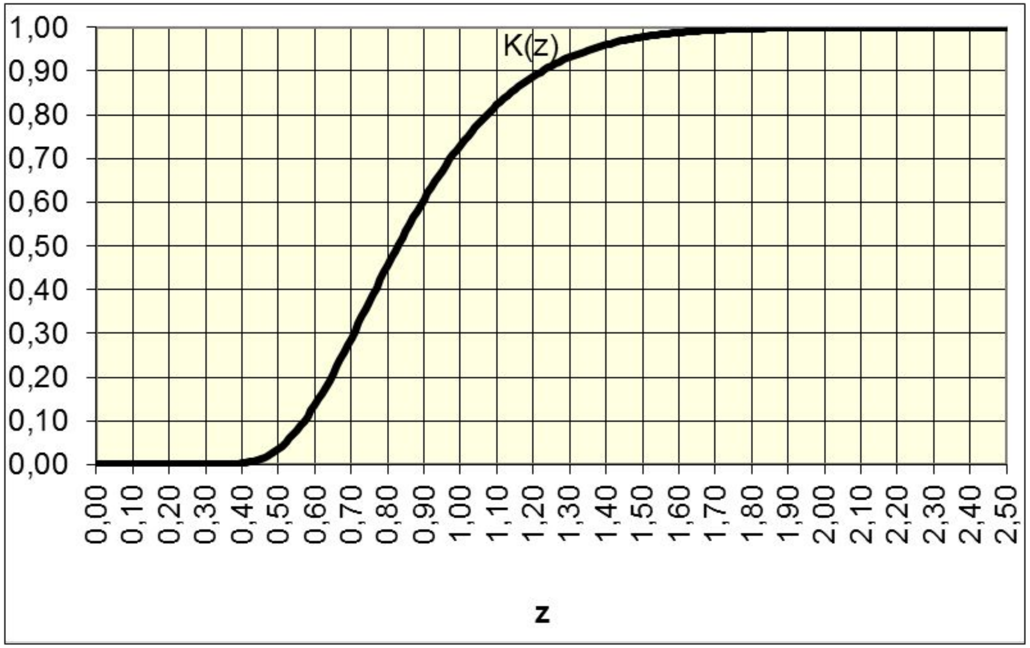
Methods of the nonstationary time series analysis

1. Ordinary least squares.
2. Time series cointegriation, i. e. the linear combination of these series becomes stationary. (Boks-Dzhenkins, 1972).
3. Autoregressive models (Dickey-Fuller, 1979).
4. Adaptive models of time series: multiparameter models of short-term forecasting(Holt,Winters, 1990-2000).

All these models operate directly with the elements of the series and predict its values. The distribution function of the series is not studied. The results depend on the length of the sample and the current time.

Nonparametric comparison of samples

Let the random variables have a continuous stationary distribution and are independent. Then the probability that the two samples of the volume N differ from each other in the norm of C by less than ε is equal to $K\left(\varepsilon\sqrt{N/2}\right)$



$$S_N = \sup_x \left| F_{1,N}(x) - F_{2,N}(x) \right|$$

$$\lim_{N\to\infty} P\left\{0 < \sqrt{\frac{N}{2}} S_N < z\right\} = K(z)$$

Distribution function as random value

If random value has distribution function F(x), than

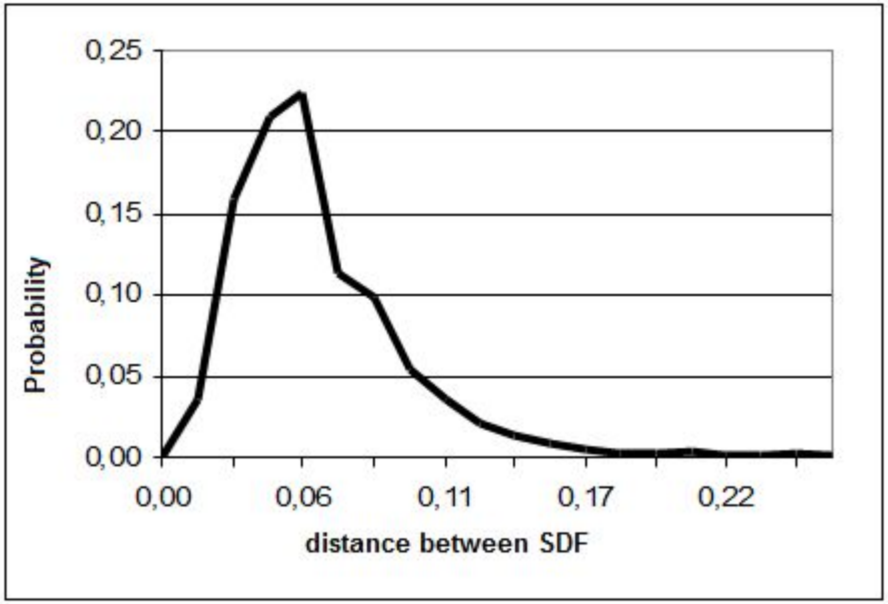
the distribution function **F**(**x**) considered as random value by itself has a uniform distribution on [0;1]

And this mean, than the quantile of uniformly distributed function is a function that depends lineary on x

Agreed level of significance (ALS)

At what distance should we unhook the "tail" of distribution of distances between distributions, so that the remaining quantile would be equal to the empirically observed level of confidence in the problem of recognizing "our" samples?

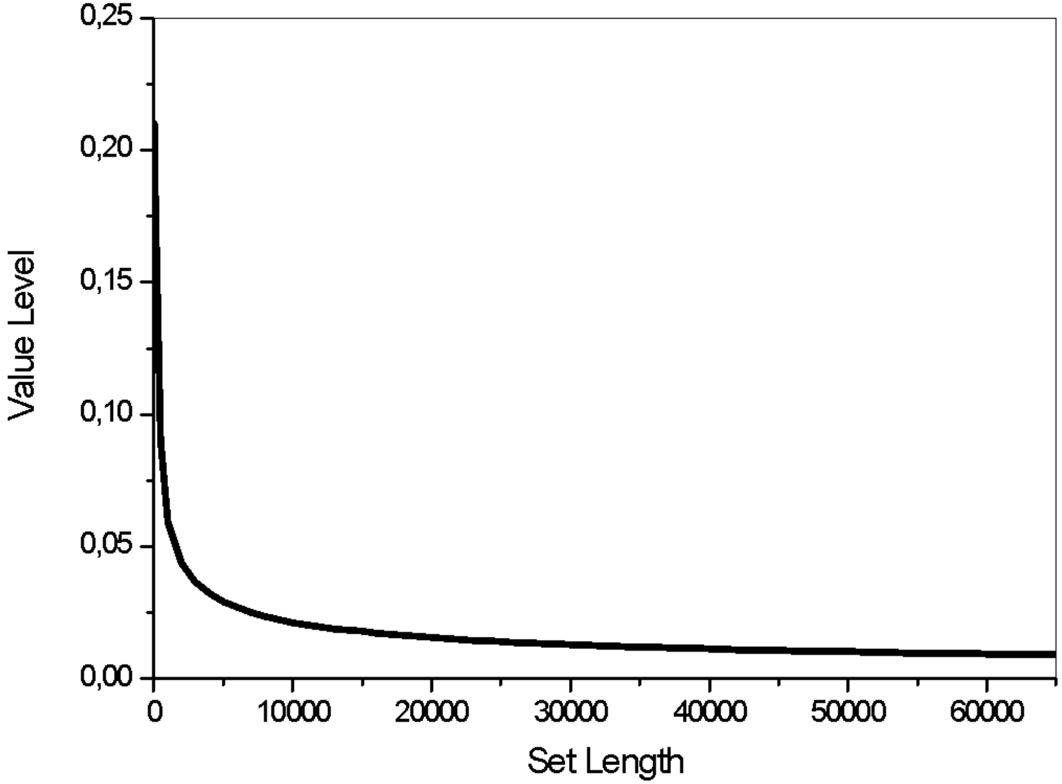
Consider the distanse **r** between the samples as random value. Its disribution function **u**=**F**(**r**) as random value is uniformly disributed. So The level of significance agreed upon with the experiment as a quantile of a uniformly distributed random value is a function that depends linearly on the distance between the samples, i.e. ɑ = ε.



In norm C, two samples of length N, the distance between which is ε, are different at the significance level α, if

$$1 - K\left(\sqrt{\frac{N}{2}}\varepsilon\right) < \alpha$$

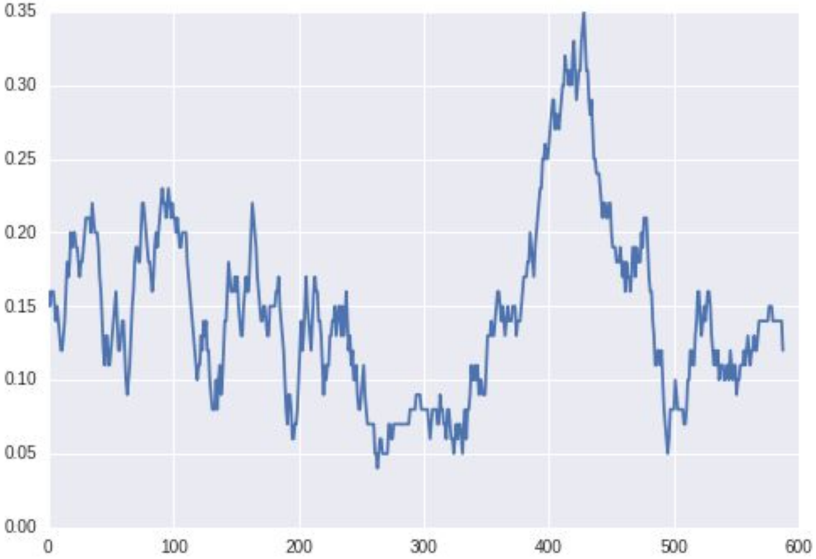
ALS in norm C

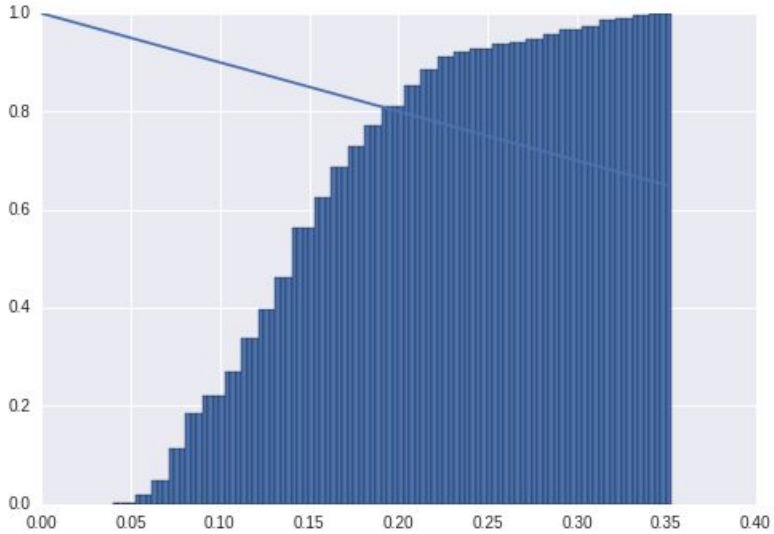


$$1 - \varepsilon = K\left(\sqrt{\frac{N}{2}}\varepsilon\right)$$

The agreed level of significance (stationarity) in the norm of C: the proportion of distances exceeding it is equal to the critical separation of samples

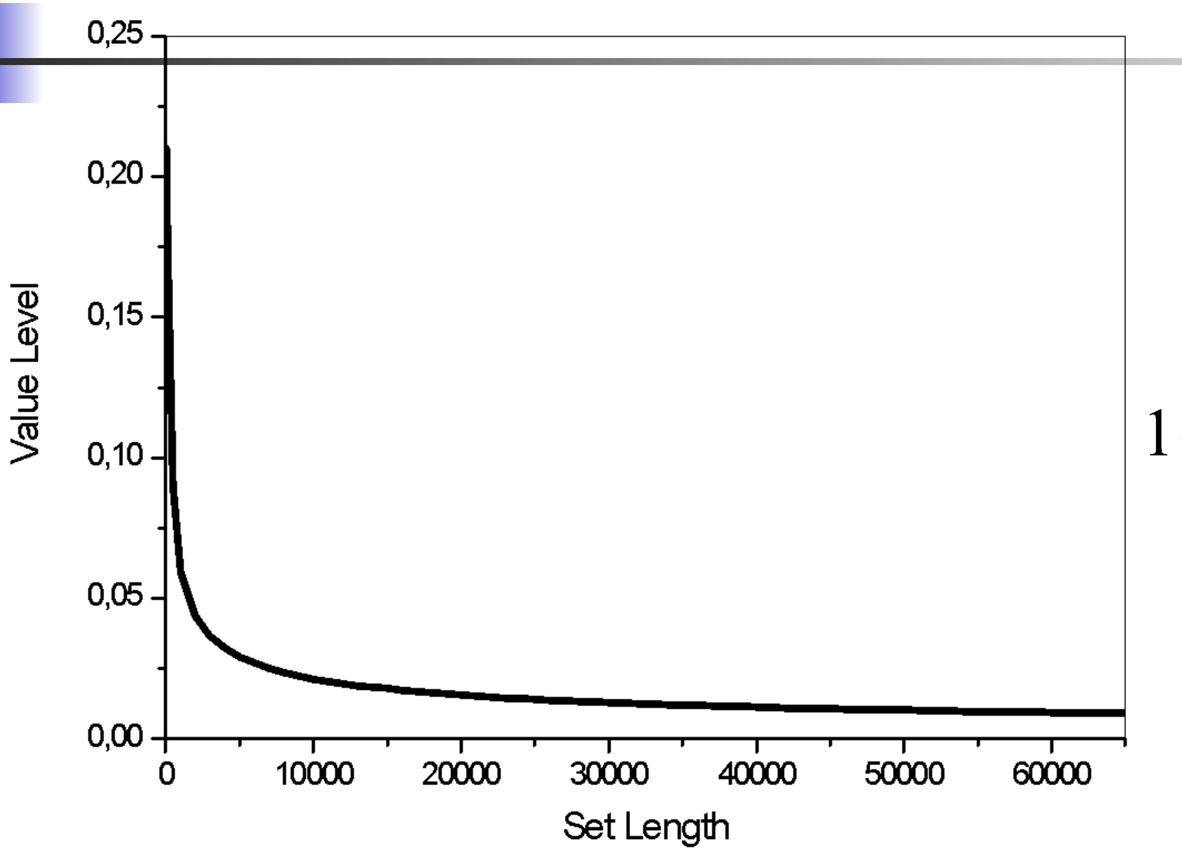
Example of ALS calculation





On the left: a series of distances between two samples of length 100 in the norm C. On the right: calculation of the ALS for the distribution of distances between distributions from samples of length 100.

Tabulation of the stationary ALS



$$1-\rho = K\left(\sqrt{\frac{N}{2}}\rho\right)$$

The ALS in the norm of C for stationary disributions does not depend on the type of distribution and is calculated from the Kolmogorov function

Nonstationary index in the norm of C

The ratio of the fraction of distances exceeding the empirical ALS is considered to the proportion of distances exceeding the agreed level of significance in the norm of C:

$$J(N) = \frac{\rho^{*}(N)}{\varepsilon(N)}$$

If J> 1, the series is nonstationary; if J <= 1, the series is stationary.

This approach allows us to introduce not the a priori, but the actual level of separation of samples in a sliding window, when the number of measurements over distributions is much larger than one.

The Fokker-Planck equation for a SDFD

$$\frac{\partial f(x,t)}{\partial t}+\frac{\partial}{\partial x}\big(u(x,t)f(x,t)\big)-\frac{\lambda(t)}{2}\frac{\partial^2 f(x,t)}{\partial x^2}=0;$$

$$u(x,t)=\frac{1}{f(x,t)}\int vF(x,v,t)dv\,,\quad v=\frac{dx}{dt};$$

$$f(x,t)=\int F(x,v,t)dv;$$

$$\lambda(t)=\sigma^2(t+1)-\sigma^2(t)-2\operatorname{cov}_{x,u}(t)=$$

$$=\frac{1}{T}\sum_{k=t-T+1}^{t}\big(x(k)-x(k+1)\big)^2-\frac{1}{T^2}\big(x(t+1)-x(t-T+1)\big)^2\geq 0$$

The sample mean and variance of the time series vary in the same way as the moments of the SDF due to the Fokker-Planck equation if the drift and diffusion are defined as written above

Method of a non-stationary trajectory generating

From the solution of the F-P equation , we know the F (x, t) at all instants of time $t_k = t_0 + k$ on the horizon N.

We generate a uniformly distributed series

$$\{y_1, y_2, \ldots, y_N\}$$

Non-stationary trajectory $\{x_1, x_2, \ldots, x_N\}$

is constructed by the inversion formula of a strictly increasing continuous function:

$$x_k = F^{-1}(y_k, t_0 + k)$$

Criteria for the correct generation of the ensemble

Let we generated s uniformly distributed rows of length N: $\{y_k\}_j, j=1,\dots,s$

Each j-th trajectory generates on the interval

$[t_0+1; t_0+N]$ SDFD $\tilde{f}_N(\{y\}_j; x, t_0+N)$, different from fact $f_N(x, t_0+N)$

Consider distances:

$$r = \left\| \tilde{F}_N\left(\{y\}; x, t_0+N\right) - F_N(x, t_0) \right\|$$

$$\rho = \left\| \tilde{F}_N(\{y\}; x, t_0+N) - F_N(x, t_0+N) \right\|$$

$$\tilde{\rho} = \left\| \tilde{F}_N(\{y\}; x, t_0+N) - \tilde{F}_N(\{y'\}; x, t_0+N) \right\|$$

ALS **r*** must be equal to the AIS of the of the original series.

ALS ρ^* must be equal to the AIS $\tilde{\rho}^*$ and both are smaller than ALS **r***.

Practical examples
Example 1 - earthquake statistics

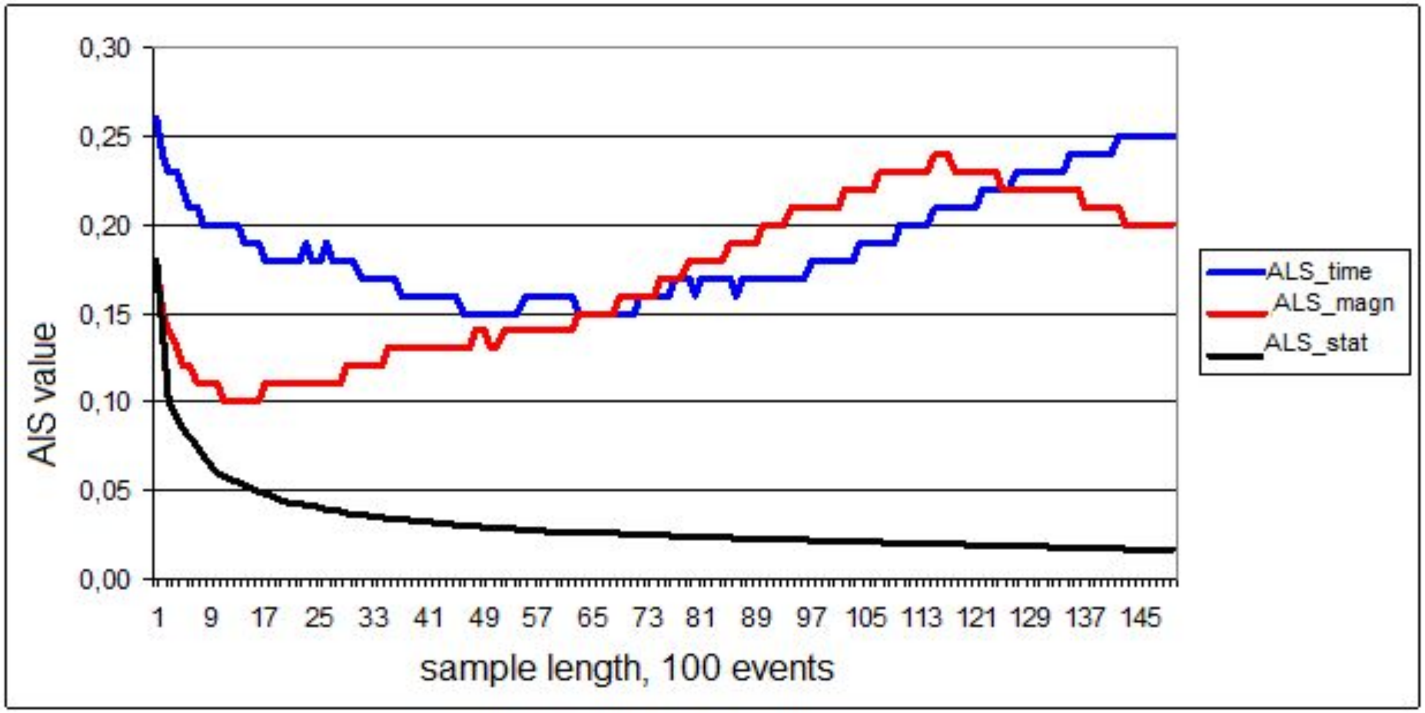
- In problems of earthquake prediction the main objects of analysis are the regional magnitudes distribution functions

  and distribution functions of time intervals between successive events. These functions shows growth or decrease of seismic activity.
- We have studied the nonstationarity of these distributions
- We have taken a time series of earthquake magnitudes in Japan from 1916 to January 2011 according to the regional catalog JMA

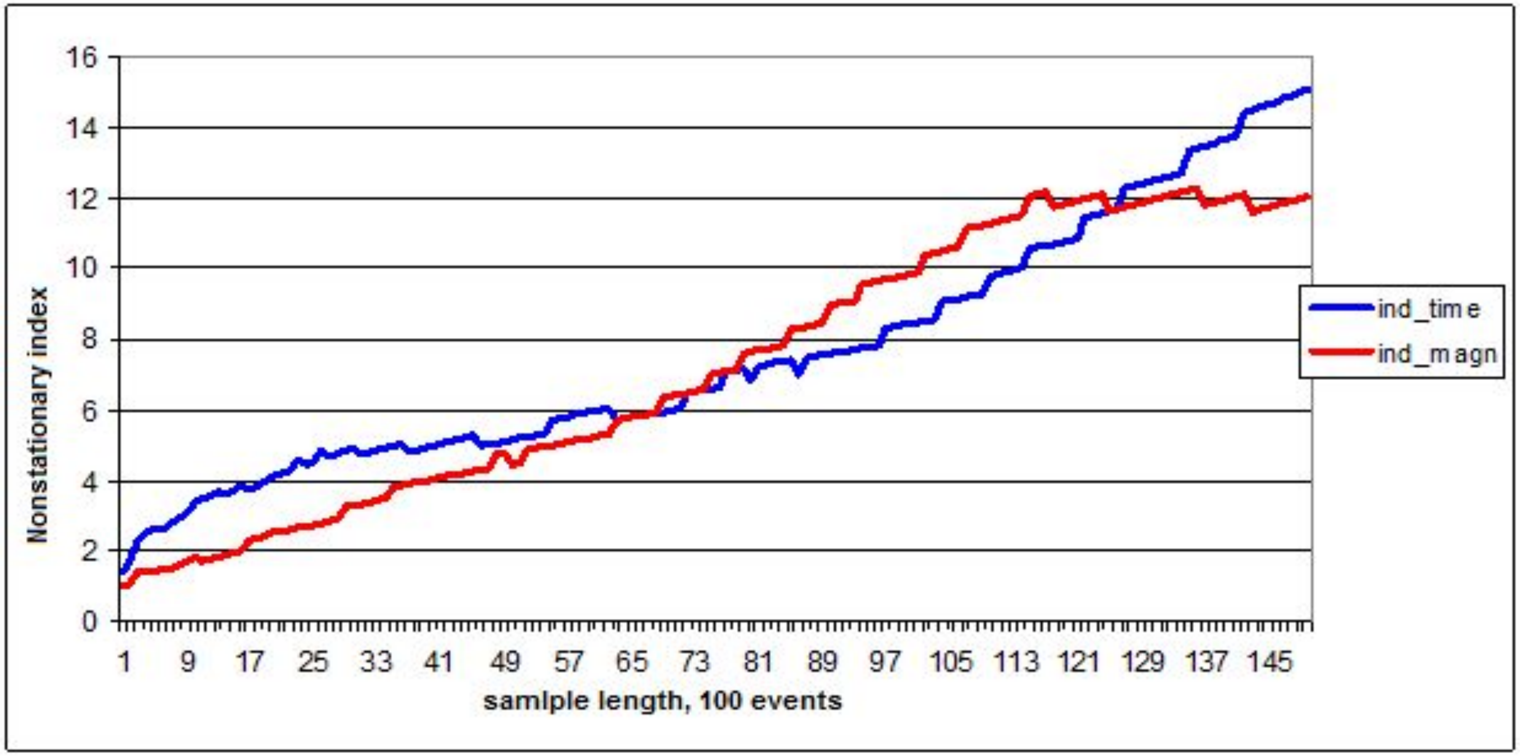
(Japan Meteorological Agency)

- Gutenberg–Richter law expresses the relationship between the magnitude and total number of earthquakes: lg(N) =a-b*M

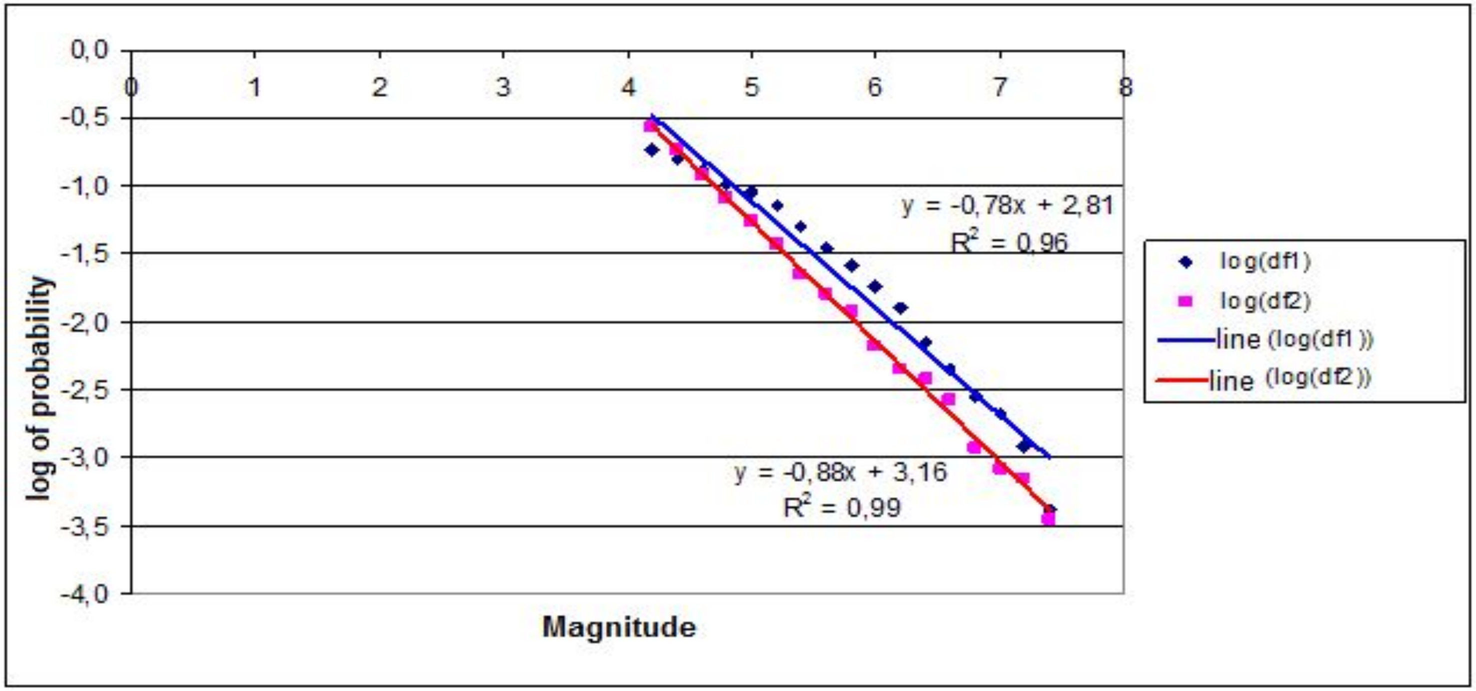
Series ALS depending on the sample length



Nonstationary index depending on the sample length

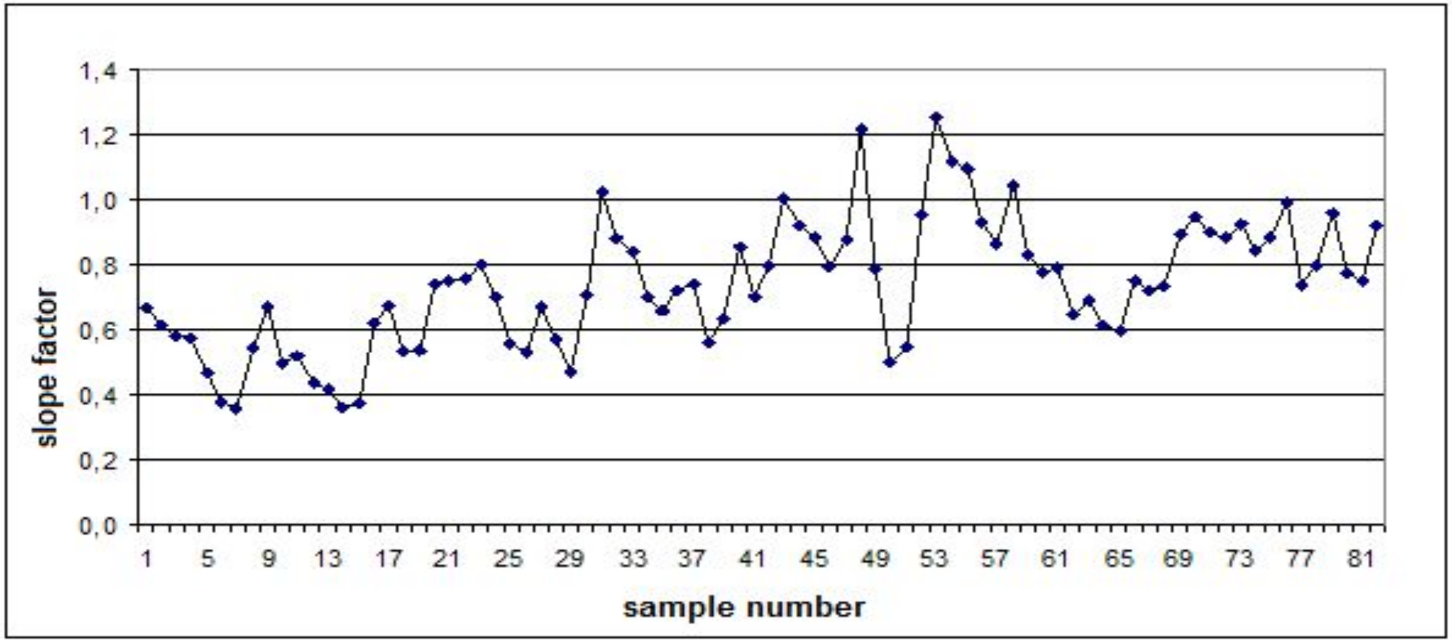


Gutenberg-Richter law for two samples



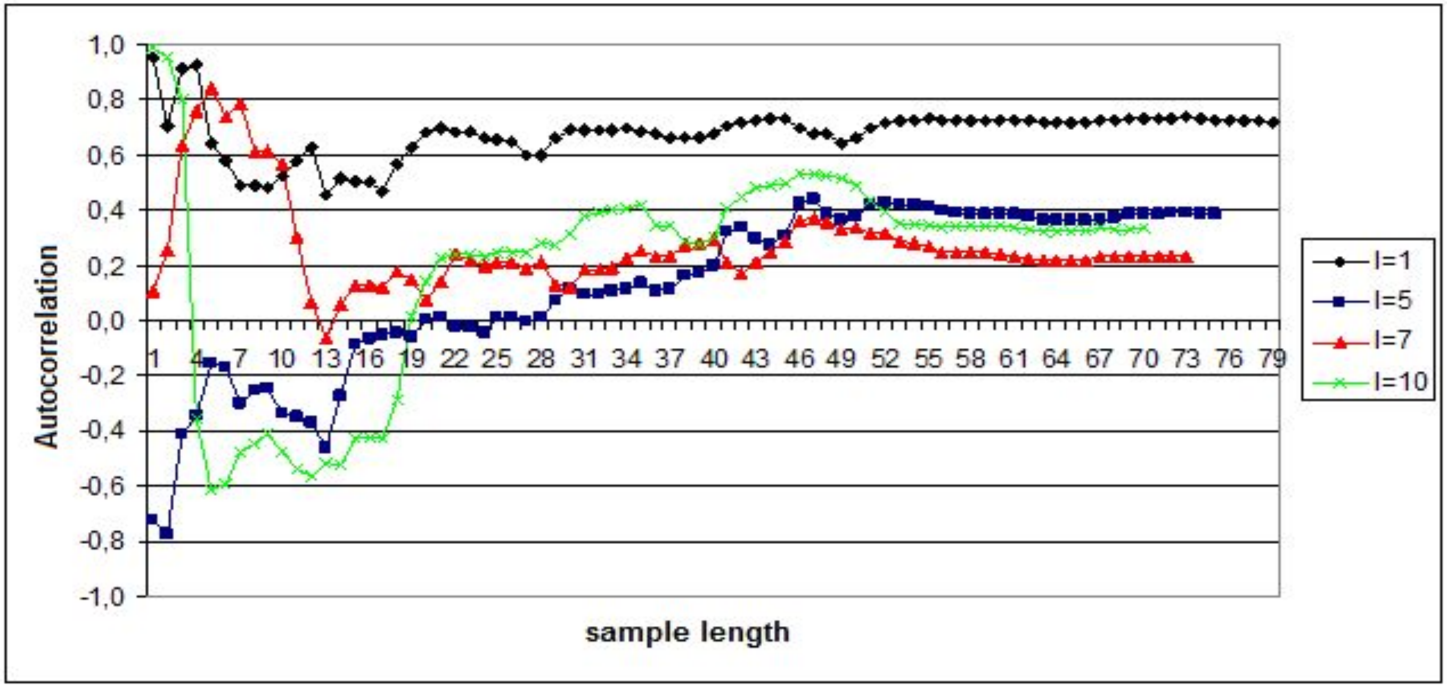
This comparison shows that the nonstationarity of the magnitude distributions can be explained by the non-stationary behavior of the slope index in the Gutenberg-Richter law, but not by the fact that the functional form of this law itself is changing.

Dynamics of the slope angle in the Gutenberg-Richter law



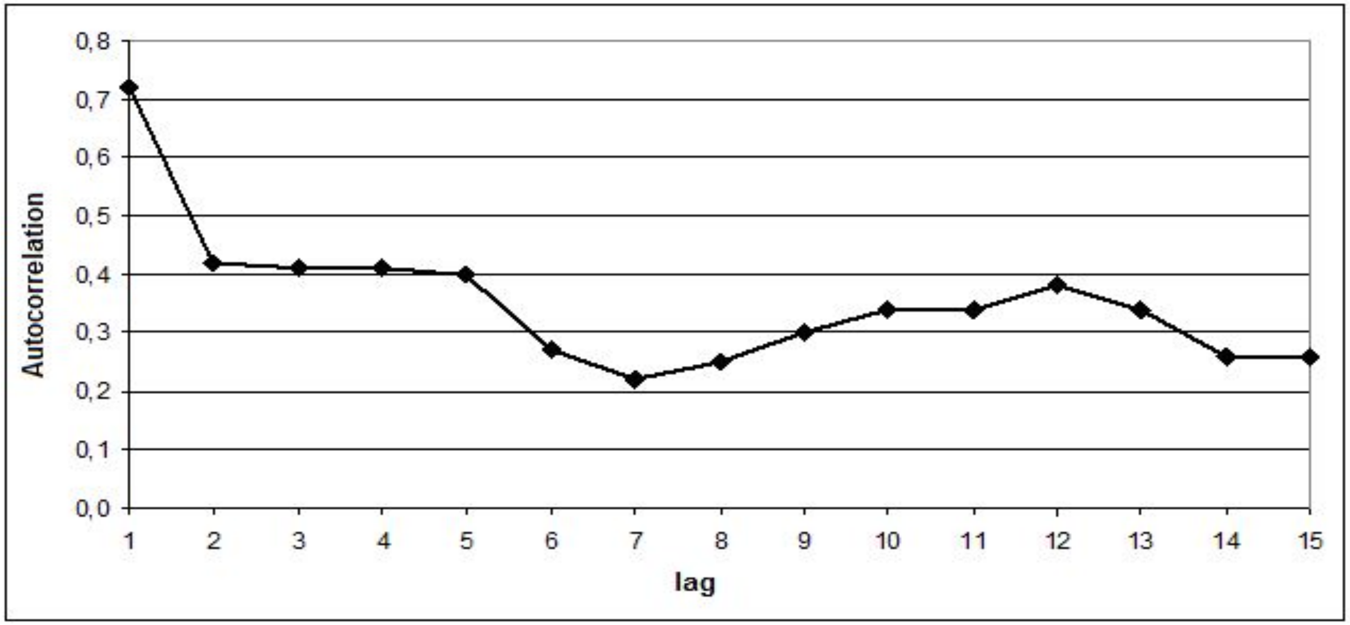
The sequence of sample slopes logarithm of the Gutenberg-Richter curve by by sliding length N=1000.

Autocorrelation analysis of the slope angle



Dependence of the autocorrelation selective coefficients of the b(n) series on the length of the sample for different lags. The values of the steady-state level are not monotone.

The values of the steady-state coefficients of autocorrelation depending on the lag



The periodicity of the autocorrelation coefficient dependence on the lag shows the presence of short-wave and long-wave quasiperiodic processes, by which the oscillatory behavior of the slope coefficient can be approximated.

The model of the time series b(n)

The dynamics of the values b(n) can be described by some quasiperiodic dynamical system with additive noise.

$$b(n) = y_1(n) + y_2(n) + \delta(n)$$

$$y_1(n) = \begin{cases} 0,45 + 0,008n, & 1 \le n \le 53; \\ y_1(53) - 0,008(n - 53), & n > 53. \end{cases}$$

$$y_2(n) = \begin{cases} 0,2 + 0,08(n - 1), & 1 \le n \le 6; \\ y_2(6) - 0,08(n - 6), & 6 < n \le 11. \end{cases}$$

Where $\delta(n)$ is a series of residues, the autocorrelation of which (for any lags) does not exceed 0.013 in absolute value, the relative mean square is 0.006, and the distribution is approximated fairly well by a normal .

Nonstationary distributions of magnitudes

it was found that the nonstationarity of the distributions
magnitude is due to the fact that the parameter in the law of the Gutenberg-
Richter depends on time, but forms a stationary time series; this
series can be represented as a superposition of two dynamical systems with
periodic behavior and a normally distributed residue that has
low amplitude

Earthquake statistics results

We analyzed the stationary level of JMA catalog of magnitude and time intervals between events. It was shown, that these distributions are nonstationary

and the time dependence of Gutenberg – Richter law parameter could be represented as a superposition of two quasi-periodical dynamical systems with short and long periods

Example 2 - SIR statistics for analysis of 5G networks

The reliability of mobile communication is estimated by the ratio of signal power to interference at the receiving point – SIR.

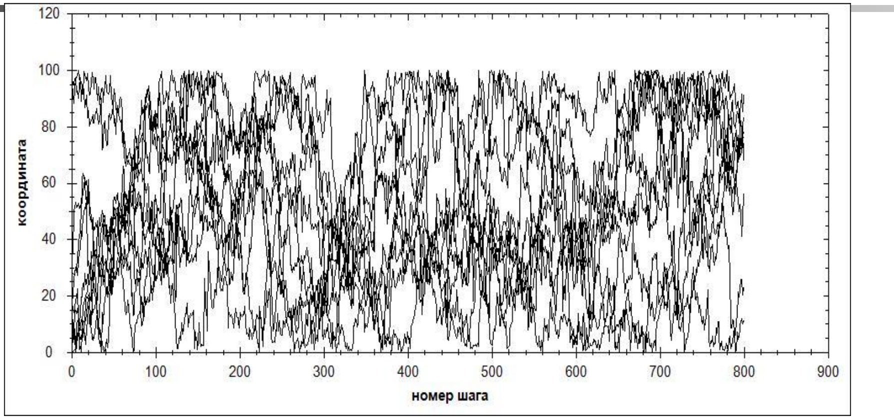
$$SIR = \frac{U(l_0)}{\sum_{i=1}^{N} U(l_i)}, \quad U(l) = l^{-\alpha}$$

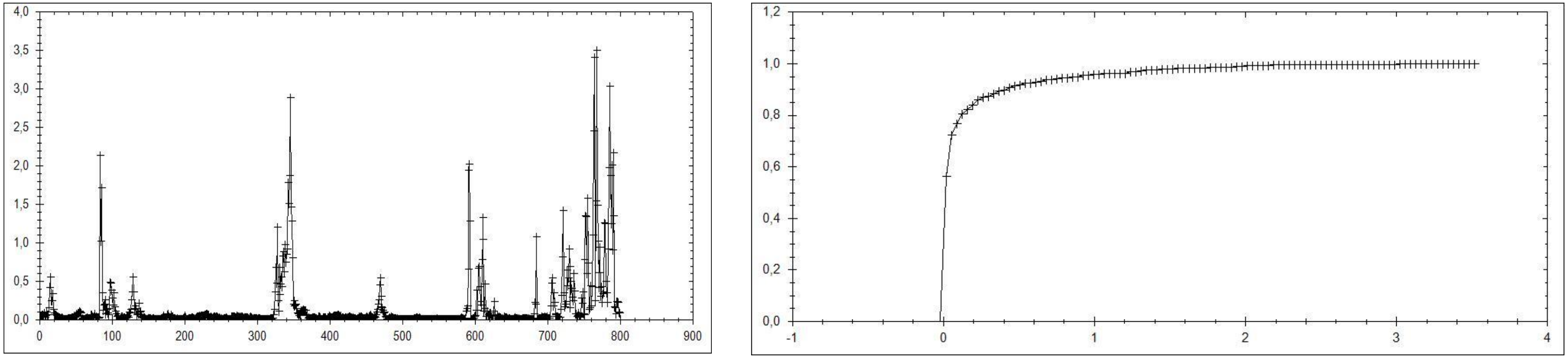
In the static mode, the SIR is analyzed by combinatorial geometry methods, but if the subscribers are in motion, then the SIR depends not only on the density of the subscribers and the shape of the region, but also on the law of motion. In many cases, the motion is stochastic and can be represented as diffusion with drift ("customer wander"). Then the trajectories of the receiving and transmitting devices are naturally modeled with the help of a suitable F-P equation:

$$\frac{\partial f}{\partial t} + div(uf) = \frac{D}{2}\Delta f$$

From the known distribution function f (x, t), a three-dimensional set of trajectories x (t) is generated, after which the distance between the corresponding points of the ensemble with confinement constraint is determined:

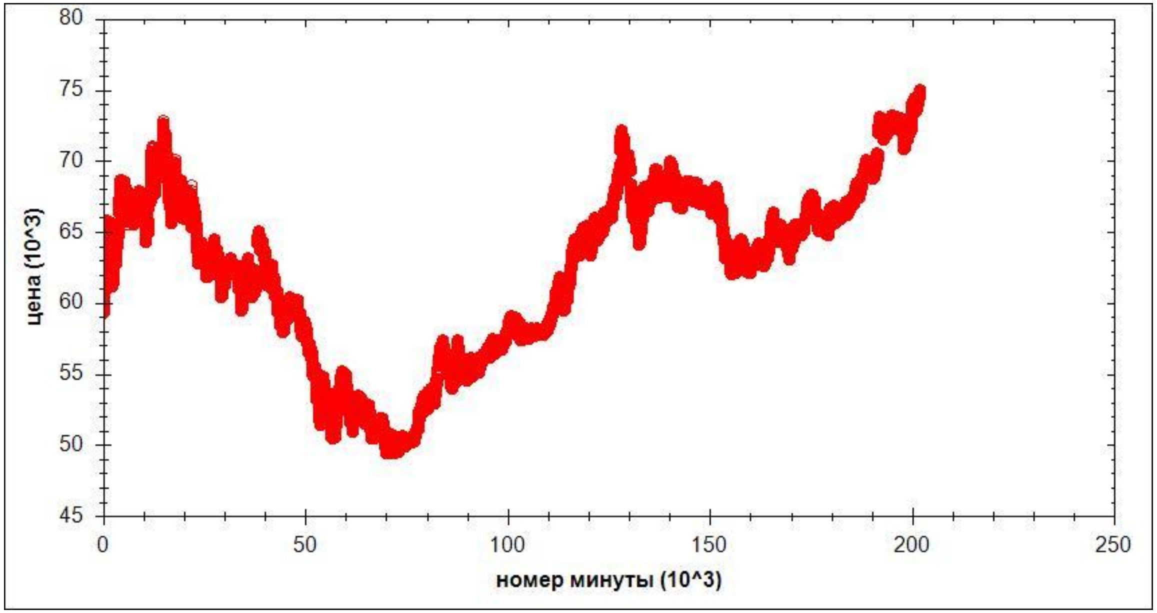
$$l_{ij}^2 = \left(x_{1,i} - x_{1,j}\right)^2 + \left(x_{2,i} - x_{2,j}\right)^2 + \left(x_{3,i} - x_{3,j}\right)^2 + a^2$$





Time series SIR (left) and DF SIR (right))

Example 3 - development of a trading system



There is a non-stationary random process (the price of the instrument) on which another process is being built - management through the functional of the trading strategy. Parameters of the strategy require testing on a long sample, which does not give a good result due to non-stationarity.

Problems types

- Selection of system parameters by historical data
- Risk-management of a trading System

Selection of system parameters by historical data

A small amount of historical data does not give sufficient accuracy. However, a large volume contains in itself not current trends.

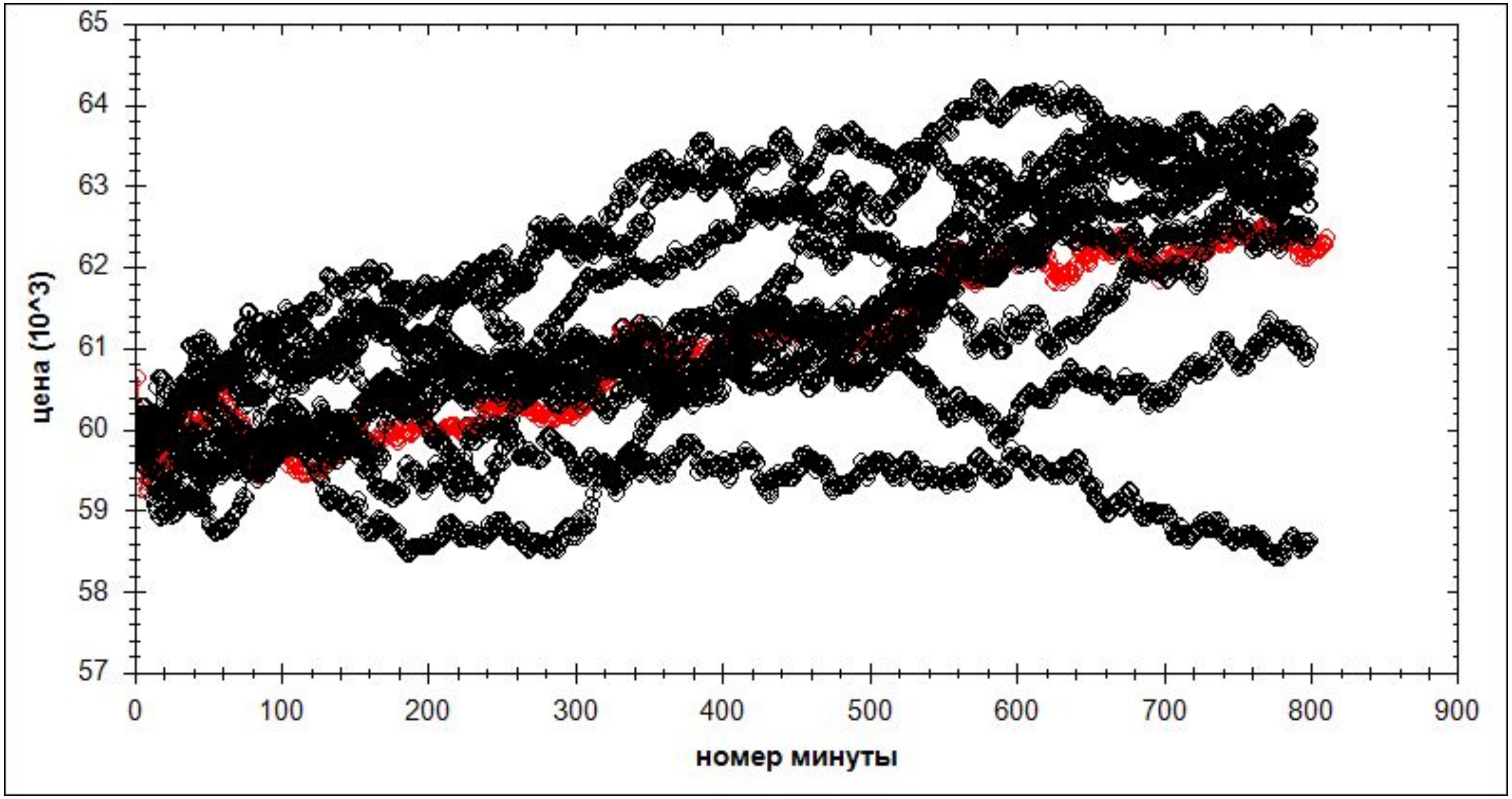
It is more efficient to generate a beam of trajectories that correspond to evolving samples in accordance with how selective distributions of price increases change.

Two types of trajectiories beam generation

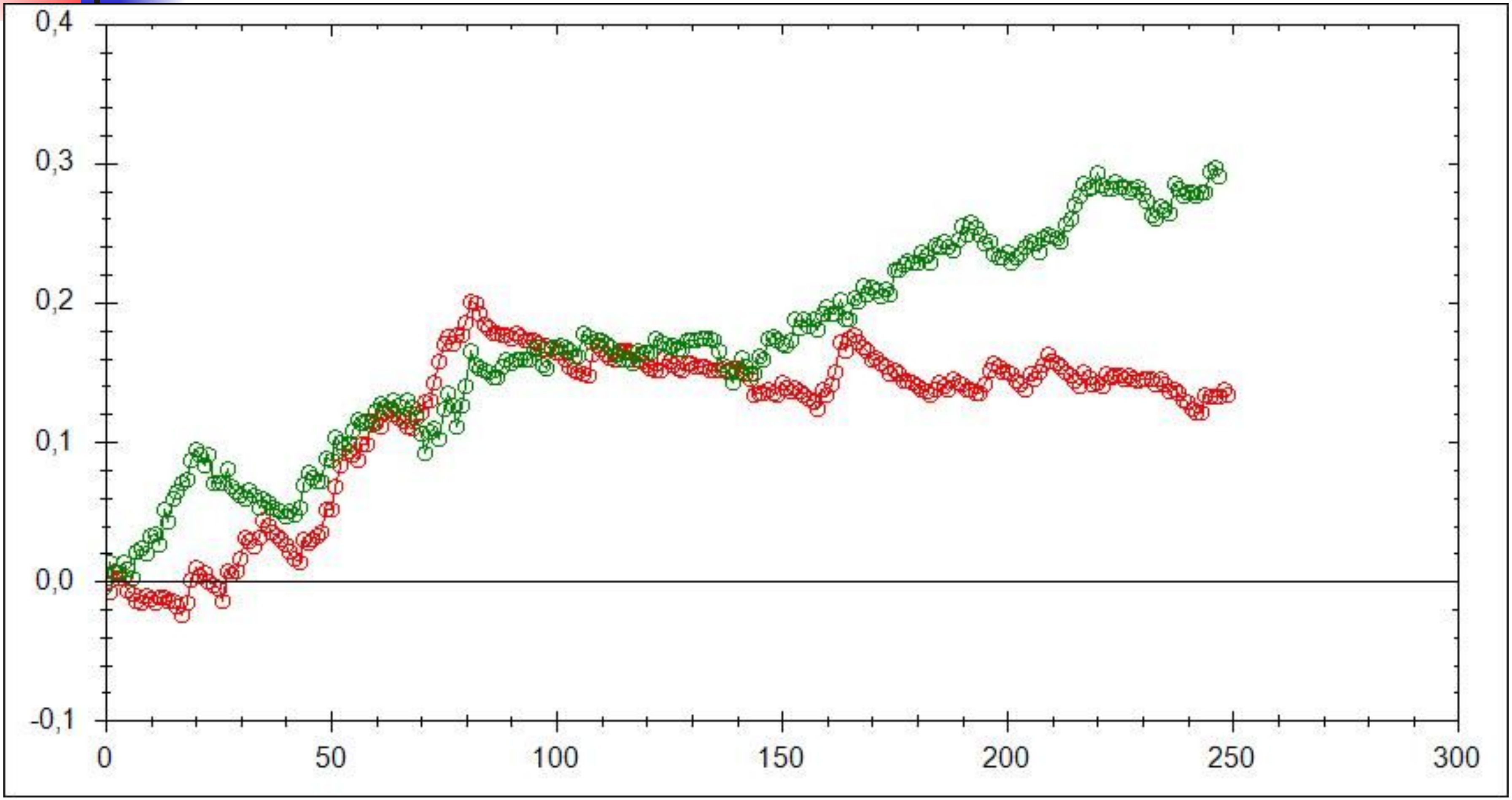
Generation by historical change of selective distribution.

Generation by forecast using Fokker-Plank equation. (Вставить формулу горизонта пересчета, когда расхождение привысит СУС)

Trajectories beam example



Strategy equities



Conclusion

Modeling of nonstationary time series has a wide practical application.

Thank you for attention!