Analysis of Caspian Sea fluctuation data

R I Dzerjinsky[1*], and *T E* Anosov[1]

[1]MIREA − Russian Technological University, 78, Vernadsky pros., Moscow, 119454, Russia

Abstract. An efficient algorithm for separating motion into fluctuating and trending components is developed. The system of almost-periods corresponding to the analyzed data is revealed using the shift function introduced by M. Johnson. Structural features of long cycles are determined, which allow building short-term, medium-term and long-term forecasts for analyzing the Caspian Sea level. It is shown that the level of the Caspian Sea in the period from the early 1950s to the early 1980s was near the minimum, and from that time began its growth. Presumably, by the early 2050s the Caspian Sea level will be able to approach the levels of the beginning of the last century and update its local maximum.

1 Introduction

The Caspian Sea is the largest enclosed body of water on Earth, which can be classified as either the largest landlocked lake or as a sea - because of its size, origin, depth, salinity, and the fact that its bed is formed by oceanic-type crust.

The Caspian Sea is subject to multi-year fluctuations in water level that can occur over several decades. These changes are usually related to natural factors such as climatic conditions, water temperature, atmospheric pressure and the influence of inter-annual climatic phenomena such as El Niño and La Niña.

There are seasonal variations in sea level at different times of the year. These fluctuations can be related to changes in wind intensity, water and air temperature, and seasonal precipitation. The history of the Caspian Sea also shows cyclical periods of rising and falling water levels. These cycles can be long-term and are related to various factors, including climate change.

North and east winds can have a significant impact on water levels in the northern part of the Caspian Sea. Strong winds can cause runup and runoff events, resulting in sea level rise or fall.

It is also worth noting that in recent decades, the impact of human activity and climate change have also become factors affecting the Caspian Sea level. Changes in the water balance, reduced water inflow from rivers and basins, and increased evaporation processes have an impact on the sea level.

Thus, Caspian Sea level fluctuations represent a complex set of factors, including natural and anthropogenic influences, which can lead to long-term changes in its hydrology.

---

[*] Corresponding author: 9015111295@mail.ru

2 Materials and methods

2.1 Data processing models and algorithms

2.1.1 Trend exclusion using the three-point method

Suppose that trend exclusion can be done by selecting the trend from the nodal points. Then, to exclude the trend, we will find their locations.

In the simplest case we can consider a situation in which three nodes $y_{t-\Delta t}, y_t, y_{t+\Delta t}$ located at equal distances relative to the central node will be used to find a solution.

In this formulation, the problem is solved by using the basic methods of the theory of proportions, according to which we divide the segment into two (Figure 1).

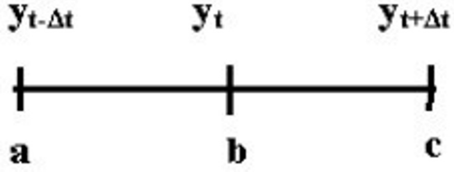

Fig. 1. Dividing a line segment.

Let us consider two kinds of proportions, arithmetic and geometric. For arithmetic proportion:

$$y_t = \frac{y_{t-\Delta t} + y_{t+\Delta t}}{2}$$

In this case, the system can be characterized as:

$$S = \frac{y_{t-\Delta t} + y_{t+\Delta t}}{2 \cdot y_t} = 1$$

Logarithmize:

$$\ln(S) = \frac{y_{t-\Delta t} + y_{t+\Delta t}}{2 \cdot y_t} = 0$$

Consider the geometric proportion:

$$y_t = \sqrt{y_{t-\Delta t} \cdot y_{t+\Delta t}}$$

In this case, the system can be characterized as:

$$P = \frac{y_{t-\Delta t} \cdot y_{t+\Delta t}}{2 \cdot y_t} = 1$$

Logarithmize:

$$\ln(P) = \frac{y_{t-\Delta t} \cdot y_{t+\Delta t}}{{y_t}^2} = 0$$

These dimensionless criteria, which represent the characteristics of the empirical series, can serve as important indicators of the properties of the system to exclude the influence of trend. The trend parameters are accounted for through the value of the shift in the argument Δt with respect to the state y_t. Thus, it is possible to transform empirical data in coordinates using these criteria.

Transformation of empirical data in coordinates:

$$\ln\left(\frac{y_{t-\tau} + y_{t+\tau}}{2 \cdot y_t}\right) \sim t \tag{1}$$

$$\ln\left(\frac{y_{t-\tau} \cdot y_{t+\tau}}{{y_t}^2}\right) \sim t \tag{2}$$

The transformation leads to the exclusion of trend areas from the data. Thus, to exclude the trend it is possible to use metrics defined by the ratio of the theory of proportions.

As the basic ones we can take a number of proportions of the Pythagorean school (Table 1).

Table 1. Types of proportions used for trend exclusion.

#	Types of proportions	Type of coordinates
1	$\sqrt{a \cdot c}$	$\ln\left(\frac{y_{t-\tau} \cdot y_{t+\tau}}{2 \cdot y_t}\right) \sim t$
2	$\frac{a+c}{2}$	$\ln\left(\frac{y_{t-\tau} + y_{t+\tau}}{2 \cdot y_t}\right) \sim t$
3	$\frac{2 \cdot a \cdot c}{a+c}$	$\ln\left(\frac{2 \cdot y_{t-\tau} \cdot y_{t+\tau}}{y_t \cdot (y_{t-\tau} + y_{t+\tau})}\right) \sim t$
4	$\frac{a^2+c^2}{a+c}$	$\ln\left(\frac{{y_{t-\tau}}^2 + {y_{t+\tau}}^2}{y_t \cdot (y_{t-\tau} + y_{t+\tau})}\right) \sim t$
5	$\frac{2 \cdot a \cdot c - a^2}{c}$	$\ln\left(\frac{2 \cdot y_{t-\tau} \cdot y_{t+\tau} - {y_{t-\tau}}^2}{2 \cdot y_t}\right) \sim t$

2.1.2 Search for almost periods

To identify periods, we can use a method that relies on the fundamental property of the period of a function, which consists in the repetition of the values of the function (3) through an interval of change of the independent variable equal to the period

$$f(t + \Delta t) - f(t) = 0 \tag{3}$$

Let us introduce the definition of the *almost-periodic function*: the number Δt is called ε-almost period of the function $f(t)$, $(-\infty < t < \infty)$ if ∇t the inequality is met:

$$f(t + \Delta t) - f(t) < \varepsilon$$

For the discrete case if n is the total number of samples of the function f(t) given the experimental values, we introduce the following metric to determine the almost periods:

$$a(\tau) = \frac{1}{n-\tau}\sum_{t=1}^{n-\tau}|f(t+\tau) - f(t)|$$

This function is called the shift function or the Alter-Johnson function.

The system of almost-periods τ of the function $f(t)$ can be defined as the set of local minima of the shift function:

$$\tau = argmin\, a(\tau), \qquad \tau_{min} \leq \tau \leq \tau_{max}.$$

Where τ_{min} and τ_{max} are the natural limits of the period search. The limits are chosen so that the value of the function $a(t)$ can be reliably determined.

The shift function is constructed as follows:

- In the first step, we take $\tau = 1$, and count the modulus of the difference of the values of the original series $|f(t+\tau) - f(t)|$, for all t from 1 to $L = N - \tau$.
- Sum the obtained values and divide by L. As a result, the first value of the function $a(t)$ has been calculated.
- Repeat steps 1 and 2 for all τ = from 2 to $N - 2$.
- The enumeration of all shifts τ allows to build the shift function.

In fact, in this case, the almost-periods presented in the initial data, independent of the fluctuation shape, are revealed.

The Alter-Johnson function reflects the structure of fluctuations of the process under study. Changes in the fluctuations of the shift function, which reflect the degree of influence of the prehistory on the formation of the current state of the process, can serve as a generalized model of fluctuations in the initial series. This property of the shift function makes it possible to use it to analyze complex fluctuation processes, including fluctuations of different durations that are not necessarily in superposition with each other.

The study of the shift function of a time series allows to solve two interrelated problems:

- To find the almost-periods of fluctuations, for the process under study.
- Determine the nature of interaction between fluctuations of different almost-periods.

The position of local minima of the shift function indicates the presence of almost-periods of corresponding durations in the initial series. If there is an almost-period τ_1 in the initial series, the shift function $a(\tau)$ will give a local minimum at the point τ_1 and, in addition, will tend to fluctuate with the same almost-period τ_1, thus forming an arithmetic progression of local minima. The fluctuations of all almost-periods present in the shift function will start from the global minimum $a(0) = 0$ at the point $\tau = 0$.

The local minima of the shift function should be analyzed sequentially, starting from the point τ=0 in the direction of increasing τ. The local minimum closest to the origin makes it possible to identify in the initial time series the almost-period τ_1 corresponding to the location of this minimum.

All further occurring local minima should be checked for multiplicity τ_1. Thus, a set of local minima will be selected, which are separated from each other by an interval equal to the first selected almost-period τ_1.

At the end of the shift function analysis, we obtain a set of independent almost-periods - τ_1, τ_1, ..., τ_n and information about the nature of interaction of fluctuations of these almost-periods with each other.

2.1.3 Model of intensive growth

In the evolution of various dynamical systems, there are not only trends or any long-lasting tendencies, but also processes that proceed rather quickly. Usually, in such processes, the reproduction rate k in equation (4) is increasing rather than decreasing. Such processes can be useful for revealing some useful characteristics of the system.

$$\frac{dy}{dt} = k \cdot y \tag{4}$$

Several models can be identified to describe such intensive growth processes:
1. The growth rate is proportional to the size of the system

$$k = a \cdot y$$

Substituting into equation (4), we obtain:

$$\frac{dy}{dt} = a \cdot y^2$$

The general equation of the process is:

$$y = -\frac{1}{a \cdot t + c}$$

The *anamorphosis* for this type of model is:

$$\frac{1}{y} = a \cdot t + c$$

This model is a hyperbola and determines the position of the vertical asymptote at which the function value reaches infinite value, which is impossible for real systems. Therefore, by the position of the vertical asymptote we can make an estimate from above to determine the time of completion of the intensive growth process.

2. The growth rate is proportional to the age of the system

$$k = a \cdot t$$

In this case it is necessary to know the beginning of the stage. Substituting k into equation (4), we obtain:

$$\frac{dy}{dt} = a \cdot t \cdot y$$

The process equation:

$$y = C \cdot e^{a\frac{(t-t_0)^2}{2}}$$

The *anamorphosis* for this type of model is:

$$\ln(y) = a \cdot \frac{(t-t_0)^2}{2} + ln(C)$$

3. The growth rate grows exponentially with the age of the system:

$$k = A \cdot e^{a \cdot t}$$

Then,

$$\frac{dy}{dt} = A \cdot e^{a \cdot t} \cdot y$$

$$y = C \cdot e^{\frac{A}{a} e^{a \cdot t}}$$

The *anamorphosis* for this type of model is:

$$ln \left[\ln \left(\frac{y}{C} \right) \right] = \ln \left(\frac{A}{a} \right) + a \cdot t$$

2.1.4 . Logistic model of constrained growth

One such model for the development of systems under resource constraints is the logistic equation:

$$\frac{dy}{dt} = k \ \cdot (y - \frac{y^2}{y_{lim}}) \tag{5}$$

Where the parameter k has the meaning of growth rate, and $\lim_{t \to \infty} y(t) = y_{lim}$ is the growth limit.

Let us determine the inflection point, which corresponds to the point of maximum production per unit time. To do this, we define the second derivative and equate it to zero:

$$\frac{d^2 y}{dt^2} = k(y_{lim} - 2 \cdot y^*) = 0$$

Where y_* is the accumulated production at the inflection point. Then$y^* = \frac{y_{lim}}{2}$, i.e. the inflection point is located by the amount of extracted resource twice below the growth limit (Figure 2). Hence, the value of the growth rate corresponding to the maximum extraction per unit of time is determined by the ratio

$$\frac{1}{y^*} \frac{dy}{dt} = k \frac{y_{lim}}{2}$$

And the maximum production per unit time will be $k \frac{y_{lim}}{4}$.

Thus, knowledge of the parameter of the logistic equation k allows to determine the position of the maximum production point and the accumulated level at the inflection point by the known growth limit.

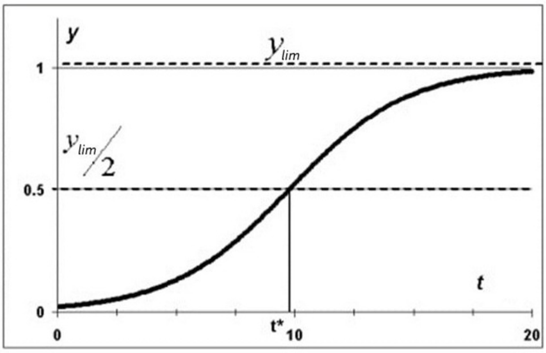

Fig. 2. Integral curve of the logistics model.

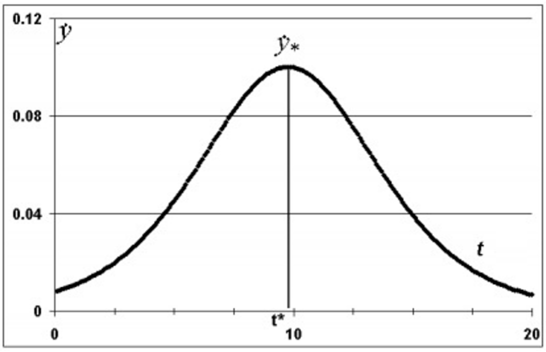

Fig. 3. Logistic model speed curve.

To determine the parameters, anamorphosis method will be used. From the original logistic model equation (5), the following anamorphosis is obtained:

$$\frac{1}{y}\frac{dy}{dt} = k(1 - \frac{y}{y_{lim}})$$

The intersection of the linear relationship in Figure 4 with the abscissa axis results in the growth limit. The parameter of the logistic model k is determined by the point of intersection of the linear dependence in Figure 4 with the ordinate axis.

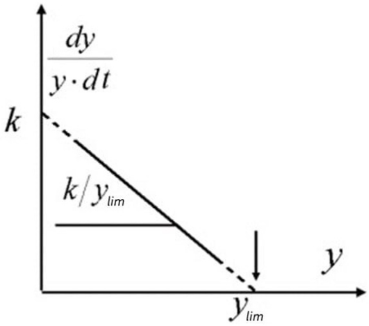

Fig. 4. Schematic of determining the growth limit of the logistic model.

2.1.5 Gompertz model of limited growth

The differential equation of the Gompertz model is:

$$\frac{dy}{dt} = A \cdot e^{-kt} \cdot y \tag{6}$$

Its integral gives $\ln(y) = -\frac{A}{k} \cdot e^{-kt} + C$. When $t \to \infty$, the constant of integration is $C = y_{lim}$. Then we note:

$$ln\left(\frac{y}{y_{lim}}\right) = -\frac{A}{k} \cdot e^{-kt}$$

Substituting the obtained ratio into equation (6), we obtain:

$$\frac{1}{y}\frac{dy}{dt} = k(lny_{lim} - lny) = -kln\left(\frac{y}{y_{lim}}\right)$$

Denoting $\frac{y}{y_{lim}} = z$ we obtain the Gompertz equation in the form:

$$\frac{dz}{dt} = -kzln(z)$$

Then the condition of maximum velocity is:

$$\frac{d^2z}{dt^2} = -k(lnz + 1) = 0$$

And at the point of maximum velocity $\ln z = -1$,

$$z = \frac{y}{y_{lim}} = \frac{1}{e}$$

Hence, the velocity reaches its peak at *e* times below the limit (Figure 5).

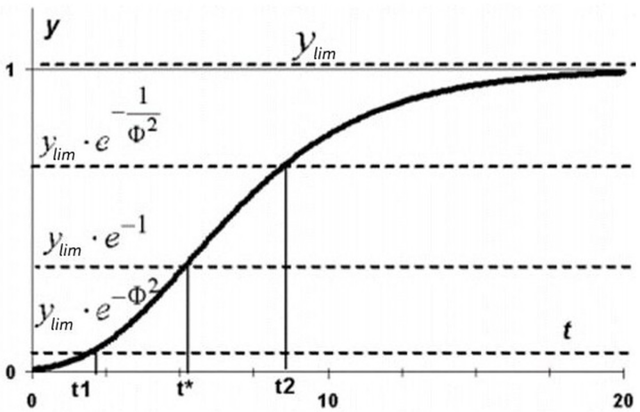

Fig. 5. Integral curve of the Gompertz model.

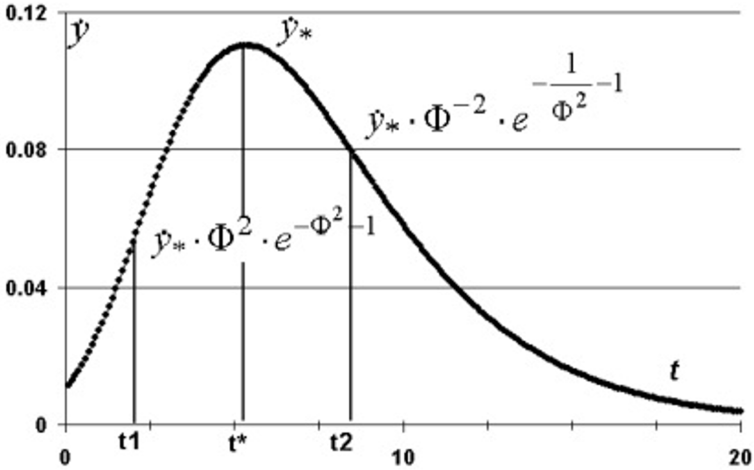

Fig. 6. Velocity curve of the Gompertz model.

From the original equation of the Gompertz model, we obtain the following anamorphosis:

$$ln\left(\frac{1}{y}\frac{dy}{dt}\right) = \ln(A) - k \cdot t$$

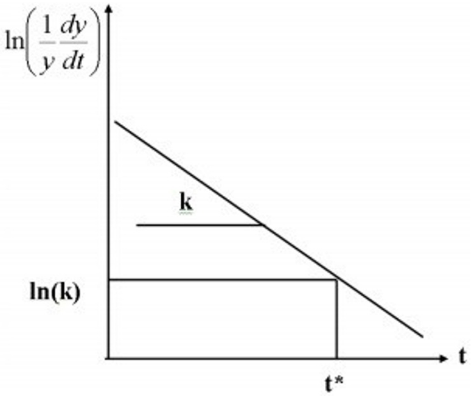

Fig. 7. Determination of the time to reach the inflection point by the value of the slope angle *k*.

3 Results and Discussion

Analysis of the Caspian Sea Fluctuation Data.

The data on the Caspian Sea level fluctuations are taken as a basis. The level of -30 meters above sea level is taken as the zero level. Let us apply the apparatus named above to study this process.

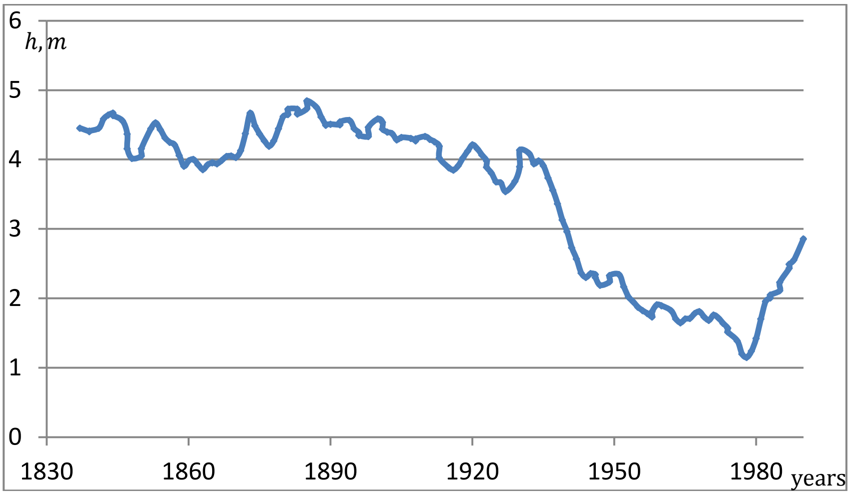

Fig. 8. Caspian Sea level dynamics.

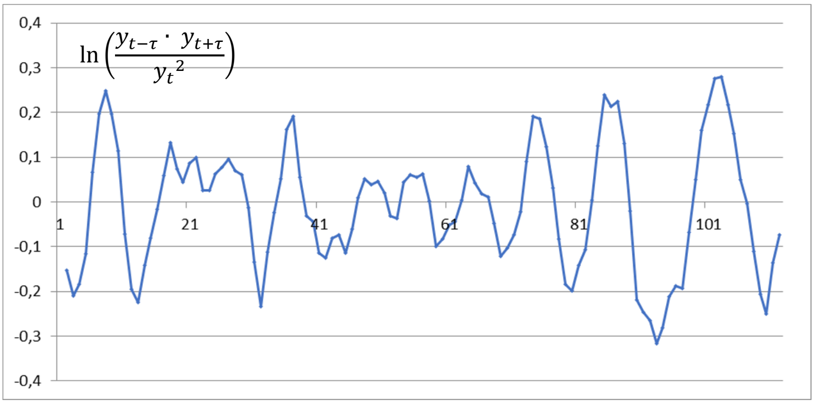

Fig. 9. Geometric trend exclusion.

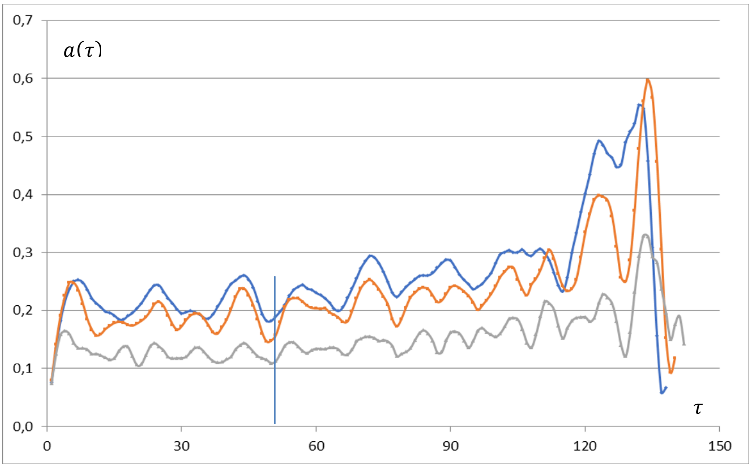

Fig. 10. Shift function.

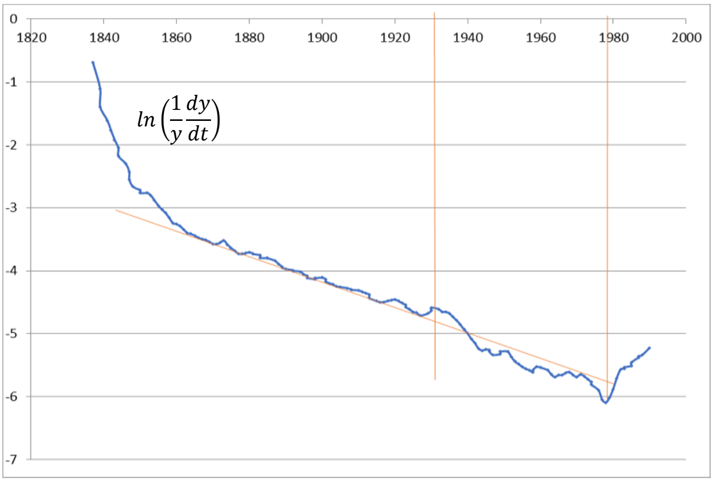

Fig. 11. Gompertz model.

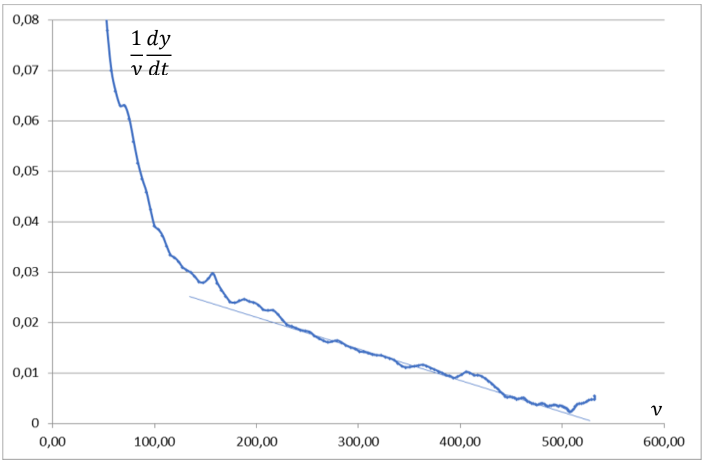

Fig. 12. Logistic model.

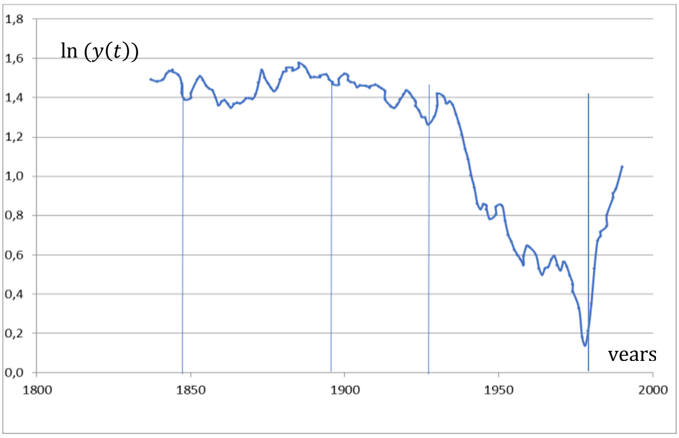

Fig. 13. Cycles in the raw data in semi-logarithmic scale.

4 Conclusion

This paper has developed a class of mathematical models and algorithms for analyzing the structure of large cycles, which allows:

- Separate empirical indicators into trend and fluctuation components in various ways.
- Carry out a comparative evaluation of the separation methods to select the optimal one.
- Obtain a set of almost-periods.
- Create predictive estimates for fast and slow motion.

In the study the methodology for the development of the long-wave theory was analyzed.

By applying the apparatus of almost periodic functions, almost-cycles of sea level change were obtained.

The influence of large solar cycles on the economy as a whole and synchronization of economic crises with changes in the Caspian Sea level, including Kondratiev cycles, were observed.

Based on the results obtained within the framework of the description of the model of intensive and limited growth, we can say that the level of the Caspian Sea at the current time is approximately in the middle of a new growth cycle that began in the late 70s - early 80s of the 20th century. In the coming years, this growth may slow down, creating a short-term sideways trend. Presumably, by the early 2050s the Caspian Sea level will be able to approach the levels of the beginning of the last century and renew its local maximum.

References

1. R.I. Dzerjinsky, S.V. Sidorov, T.E. Anosov, The Problem of Large Local Fluctuations Appearance, Lecture Notes in Networks and Systems, **723**, 573–580 (2023)
2. N.D. Kondratiev, Y.V. Yakovets, L.I. Abalkin, Big cycles of conjuncture and the theory of foresight. Selected Works, Ekonomika, Moscow (2002)
3. V.I. Kuzmin, A.F. Gadzaov, Mathematical methods of periodic components of nonlinear processes and predict the dynamics of the limited growth based on them. Russian Technological Journal, **4-2**, **9**, 94-104 (2015)
4. A.V. Korotayev, The Laws of History: Century Cycles and Millennial Trends. Demography, economy, wars, KomKniga, Moscow (2007)
5. R.K. Klige, “Changes in the water regime of the Caspian Sea,” GeoJournal, July, 299-307 (1992)
6. S.Yu. Glazyev, Modern theory of long waves in the development of the economy. Economic Science of Modern Russia, **2**, **57**, 8-27 (2012)
7. A.N. Kolmogorov, S.V. Fomin, Elements of the theory of functions and functional analysis, Fizmatlit, Moscow (2004)
8. A.D. Myshkis, Elements of the theory of mathematical models KomKniga, Moscow (2007)
9. V.A. Sadovnichiy, A.A. Akaev, A.V. Korotayev, S.Y. Malkov, Modeling and Forecasting of World Dynamics, ISPI RAN, Moscow (2012)