Analysis of Primary Energy Consumption in the United States

A. A. VYSOTSKAYA, A. A. VORONTSOV, R. I. DZERZHINSKY
Applied Mathematics Department
MIREA - Russian Technological University
78, Vernadsky Ave., Moscow, 119454
RUSSIAN FEDERATION

Abstract: This article examines the overall trend in primary energy consumption from 1635 to 2020 in the United States. Based on the exponential growth model, the annual energy consumption average growth rate and the process characteristic time are determined. The anamorphosis method is used to search for the logistic model parameters. Long-term trends analysis and the corresponding time series fluctuations are carried out. For the fluctuations' analysis, the trend component is preliminarily excluded based on the proportions theory. Near-period values are determined using shift and autocorrelation functions. To predict further energy consumption dynamics, the ARIMA autoregressive model is used, on which basis a local increase in the annual energy consumption level to 97.66 quads Btu is expected by 2025. The US energy consumption dynamics by resource type is considered. A forecast up to 2025 for the primary energy resources consumption shares in their total volume is built on an autoregressive model basis.

Keywords: analysis, time series, energy consumption, autocorrelation function, shift function, trend, development cell, forecast, autoregressive model

Received: February 16, 2021. Revised: August 5, 2021. Accepted: August 22, 2021. Published: August 31, 2021.

1 Introduction

The energy market constitutes a country's economy a large sector, so different development forecasting methods are being developed for it. The level of development of each country and the world economy as a whole is tightly linked to energy consumption. First, the prospects for the energy market development depend on the state economic and energy plans strategies, which directly affect each other. All mechanisms for solving problems facing the energy market are based on energy saving, energy efficiency and energy substitution technologies. In addition to economic consequences, the global energy problem is closely related to the environment and the negative environmental consequences of the energy production and consumption processes, which lead to climate change and global temperature rise. Therefore, the key issue is the energy problem solution, which requires the countries remaining energy resources accurate assessments, the search for new energy sources, the new technologies' creation for energy resources extraction. Studying the dynamics and obtaining forecast estimates of the primary energy resources consumption in the United States seems to be relevant, since it can make certain adjustments to the energy future world production and consumption forecast estimates, both at the global level and at the regional level.

Such a task leads to the need to study the energy market development processes interaction representations using fundamentally different time scales using the near-periodic functions apparatus, which is achieved by dividing the initial time series into components that characterize different-speed ones, conventionally defined further as fast and slow process components. The rapid movements' analysis allows the time intervals' coherence closest to near-periods to be determined.

Knowing the near-periods allows us to use them as the initial series smoothing intervals to determine the trends' hierarchy corresponding to the near-periods obtained values and to identify the so-called development cell that determines the prognostic observation period.

Near-proportions, in this case, can be detected both using algorithms for finding almost-proportions and can be reproduced through near-period ratios [1, 2, 3].

The consideration' such an ideology application makes it possible to analyse different nature markets [4, 5].

This paper looks at the movement' oil consumption from 1635 to 2020 in the United States [6].

2 Exponential Growth Models

Consider the annual primary energy consumption dynamics, obtained on the Energy Information Administration (EIA) regular statistical reviews basis [6]. The data collected by the EIA in quads Btu (Btu - British thermal unit, 1 kWh = 3412 Btu, 1 quadrillion = 10^{15}) units are as complete as possible since it takes into account primary energy resources' all types: coal, gas, liquid fuels, renewable energy sources (geothermal, solar, wind, etc.), nuclear energy. Figure 1 shows the data in a Cartesian coordinate system, in Fig. 2 - in a semi-logarithmic coordinate system, which is a natural way to identify areas of an exponential trend, which is characterized by the relative increments constancy.

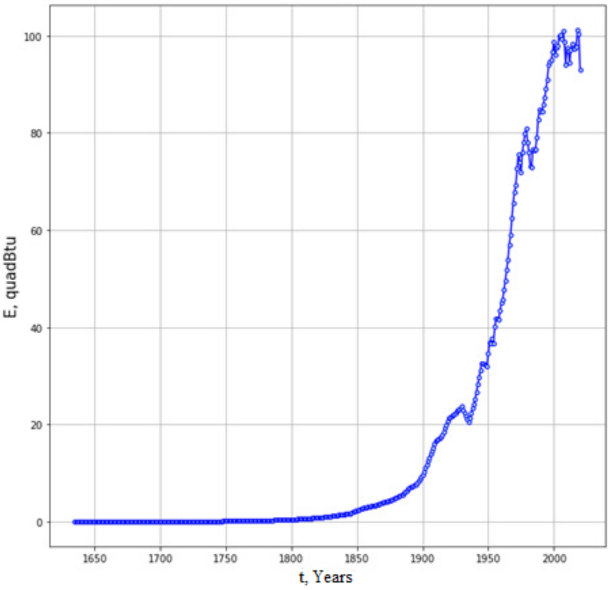

Fig. 1. Primary energy resources consumption dynamics in the USA, 1635-2020.

The system growth rate is calculated by the formula:

$$r(t) = \frac{\ln\frac{E(t_2)}{E(t_1)}}{t_2 - t_1}, \quad (1)$$

where: *t* is the year,

r(t) - growth rate on the interval [*t*, *t*+1],

E(t) - energy consumption during the year *t*.

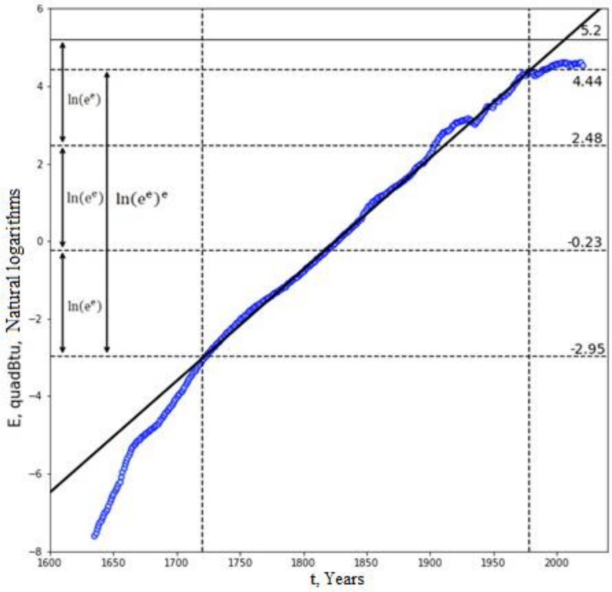

Fig. 2. Semi-logarithmic coordinate system.

Also, this value is determined by the linear dependence inclination angle in semilogarithmic coordinates and its meaning corresponds to the process inverse characteristic time in the linear section.

In the period 1720-1978, the annual energy consumption growth rate remains constant at about 2.9% per year. The process characteristic time is 34.5 years. Every 34.5 years, the annual energy consumption level increases 2.72 times. In the initial coordinates, this development area corresponds to an exponential regime,

$$E_{exp}(t) = E_0 \cdot e^{r(t-1720)}. \quad (2)$$

The US economy development period from 1720 to 1978 is the development main mode [7], which has a balanced character and maximum growth rates. After 1978, the growth in annual energy consumption ceases to follow an exponential law. Further, there is a departure from the highway, which is accompanied by a decrease in growth rates. Fig. 2 also shows a drop in relative growth, which is typical for limited growth processes.

3 Problem Solution

On the limited growth models basis, it is possible to effectively predict the system growth redistribution. One of the most common limited growth models is the logistic model. The logistic differential equation looks like this:

$$\frac{dE}{dt} = k \cdot E \cdot (E_\infty - E). \quad (3)$$

The solution to this equation is the logistic curve:

$$E_l(t) = \frac{E_\infty}{1+e^{-k\cdot(t-t^*)}}. \quad (4)$$

The only way to efficiently identify a model parameter is to plot empirical data in coordinates that give straightening. For this, anamorphosis is used - nonlinear transformations that bring data into a linear relationship. Anamorphosis for the logistic model:

$$\frac{dE}{Edt} = k(E_{\infty}\text{-}E). \quad (5)$$

Straightening is achieved in coordinates $\dot{E}/E \sim E$. Let's build a graph in these coordinates (see Fig. 3). The trend line intersection with the abscissa axis defines the position of E_{∞}, and the intersection with the ordinate axis defines the model parameter *k*.

We get the model parameters: $E_{\infty} = 12169.5$, $k = 0.0331$. The process characteristic time is 30.2 years. The inflexion point $t^* = 2013$ is defined as the point in time at which the accumulated energy consumption level is equal to half of the found growth limit. Under the logistics law, in 2013, the annual energy consumption growth trend ended.

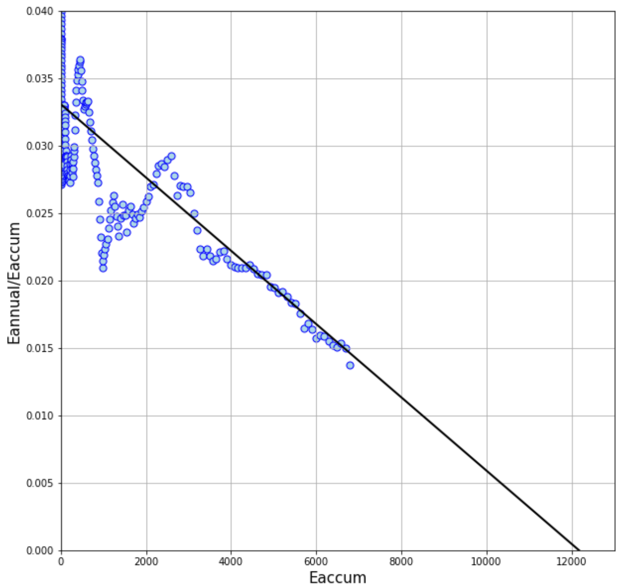

Fig. 3. Anamorphosis for the logistic model.

From the known growth limit E_{∞}, it is possible to determine the annual energy consumption maximum level,

$$\sum_{\infty} \cdot \frac{K}{4} \approx 100.7 \text{ quads Btu} \quad (6)$$

The resulting estimate is in line with the energy consumption actual level achieved in 2007. Consequently, annual US primary energy consumption has now reached a 'plateau' and is in the high point vicinity, as shown in Table 1.

Table 1. Primary energy consumption level over the last 13 years.

Year	E, quads Btu
2008	98.75
2009	93.94
2010	97.51
2011	96.86
2012	94.37
2013	97.12
2014	98.28
2015	97.37
2016	97.34
2017	97.59
2018	101.16
2019	100.29
2020	92.97

4 The US Energy Consumption Dynamics Analysis by Resource Type

Consider the consumption data for individual energy resources, the sum of which represents total US energy consumption. Fig. 4 shows the energy consumption dynamics from liquid fuels, gas, coal, renewable sources and energy from nuclear sources from 1949 to 2020.

The United States consumes its energy most from oil, followed by gas. There is also an increase in energy consumption from renewable sources at an increasing pace.

In Fig. 5, solid curves show the energy resources various types annual consumption shares in their total volume, dashed curves show a forecast based on the ARIMA model for resource' each type for the next 5 years.

There is a declining trend in the energy consumption share from oil and coal in the United States, while the energy share derived from gas is increasing.

Forecast estimates are presented in Table 2.

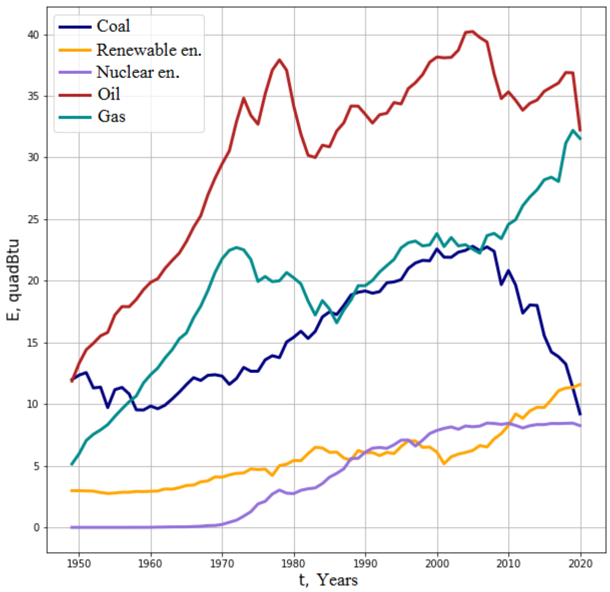

Fig. 4. Energy consumption dynamics.

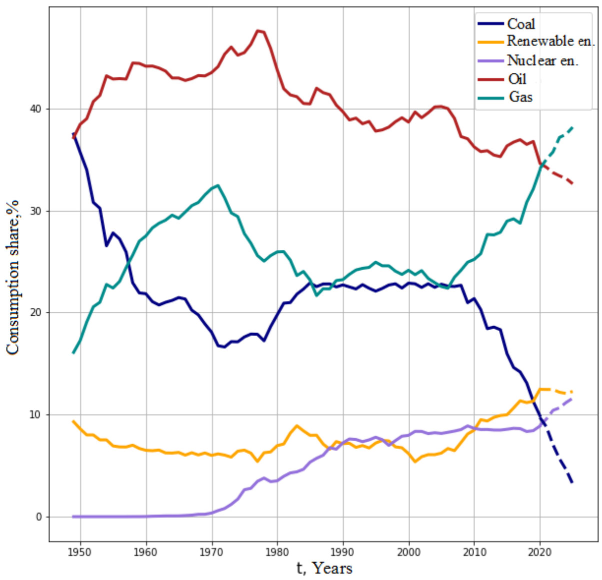

Fig. 5. Consumption shares and forecast.

Table 2. The energy consumption share forecast estimates by resources types.

Resource	Consumption percentage per year				
	2021	2022	2023	2024	2025
Oil	34.22	33.7	33.38	33.12	32.61
Coal	8.80	7.05	5.67	4.60	3.22
Gas	35.00	35.72	37.15	37.45	38.15
Nuclear energy	9.56	10.41	10.66	11.18	11.59
Renewable sources	12.46	12.46	12.19	12.07	12.27

5 The Structural Structure Patterns Analysis Manifested in the Energy Consumption Dynamics

It is quite problematic to analyse fluctuations since the initial data contain a trend component. It distorts the result. Therefore, it is first necessary to exclude the trend.

The trend characteristics are considered to be encoded through the anchor points. The task is to find such these points a position, which will ensure the trend exclusion. The simplest case is when only three points $y_{t-\Delta t}, y_t$ and $y_{t+\Delta t}$ are used for the solution, which is at an equal distance from each other. In this formulation, the problem is reduced to the proportions' theory classical results [2], according to which the segment is partitioned.

Convert empirical data to coordinates:

$$\ln\left(\frac{y_{t-\Delta t}+y_{t+\Delta t}}{2\cdot y_t}\right)\sim t, \qquad (7)$$

$$\ln\left(\frac{y_{t-\Delta t}\cdot y_{t+\Delta t}}{{y_t}^2}\right)\sim t, \qquad (8)$$

leads to the trend areas exclusion from the data.

Applying the transformation (8) to the primary resources' total energy consumption data in the United States with various shifts Δt, we obtained the result that is shown in Fig. 6.

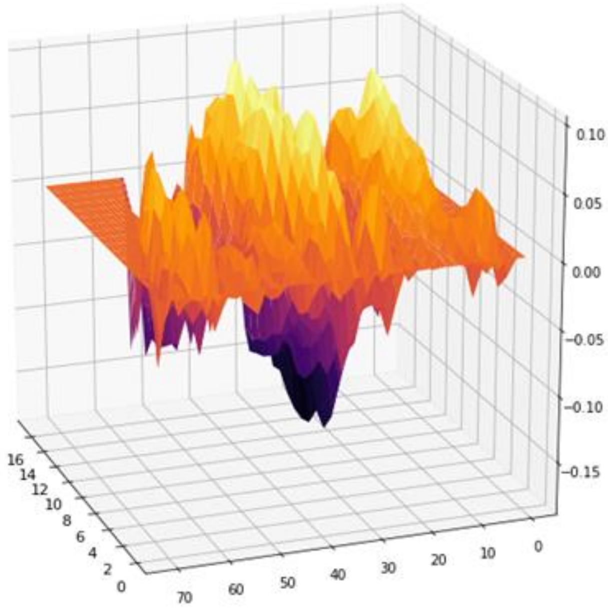

Fig. 6. Excluding the trend result.

When analyzing time series, it is necessary to know the fluctuations near-periods. These are the values that are closest to the period true value. To identify them, the shift function or the Alter-Johnson function is used, and it looks like this:

$$a(\tau)=\frac{1}{n-\tau}\sum_{t=1}^{n-\tau}\left|f(t+\tau)-f(t)\right|, \qquad (9)$$

where:

n is the function *f(t)* samples total number,

f(t) is a periodic function,

τ - function period.

This approach is based on a function period fundamental characteristic property, which is that the function values are repeated at variable equal variation an interval to the period. The shift function local minima position indicates the different durations near-periods presence in the initial time series.

Another function that can be used to determine a system near-periods is the autocorrelation function or ACF. The autocorrelation function allows finding time series repeating portions or to determine the carrier frequency, which may be hidden due to the noise or fluctuations superposition at other frequencies. When ACF is applied, repeating areas are identified by overlaying the series on itself, but with different shifts τ in time.

The autocorrelation function looks like this:

$$w(\tau)=\frac{1}{n-\tau}\sum_{t=1}^{n-\tau}f(t+\tau)\cdot f(t). \qquad (10)$$

The function local maxima determine the near-periods position.

Let's apply the shift and autocorrelation functions to the data without a trend. The shear function is shown in Fig. 7.

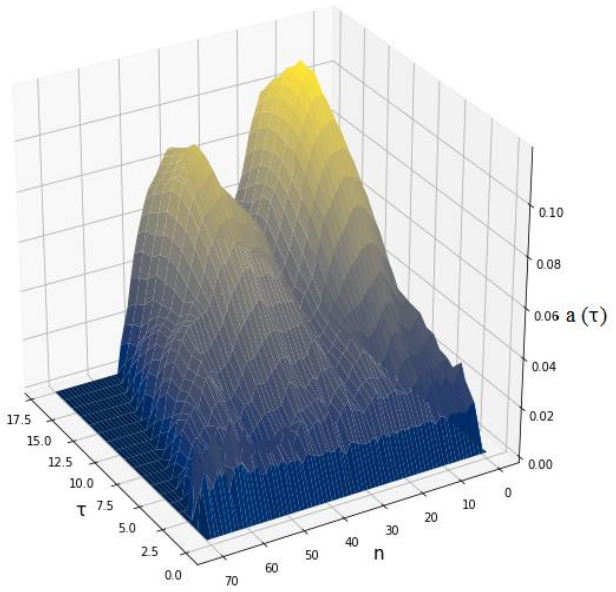

Fig. 7. Shear function.

The shear function level map is shown in Fig. 8.

On the shear function levels map, three rhythms are distinguished: 3-4 years, a multiple of 7-8 years, and a triple of 10 years.

The autocorrelation function is shown in Fig. 9.

The autocorrelation function level map is shown in Fig. 10.

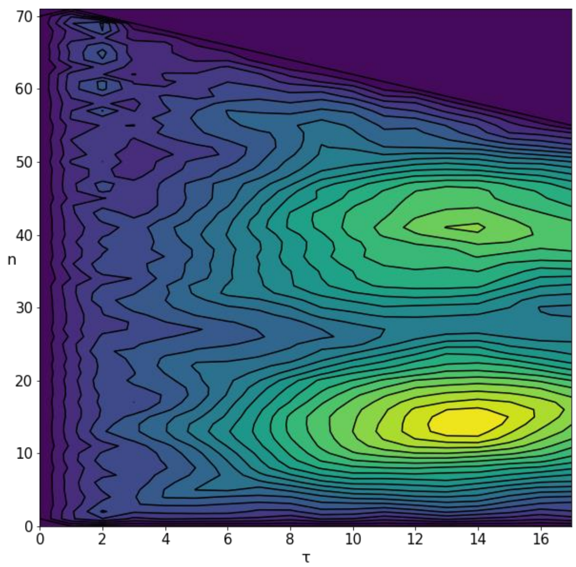

Fig. 8. The shear function levels map.

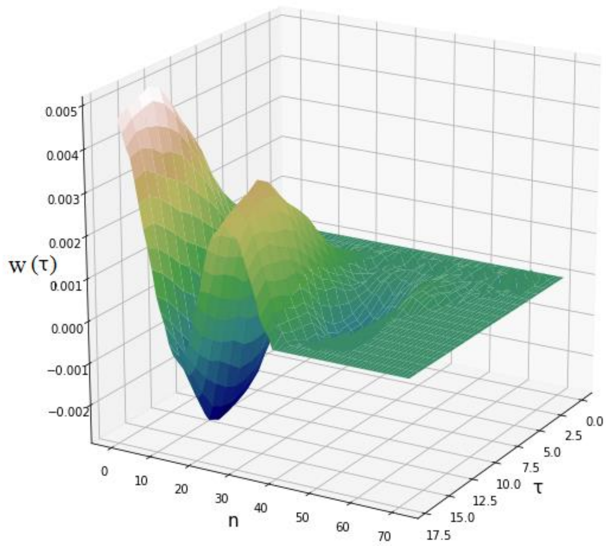

Fig. 9. Autocorrelation function.

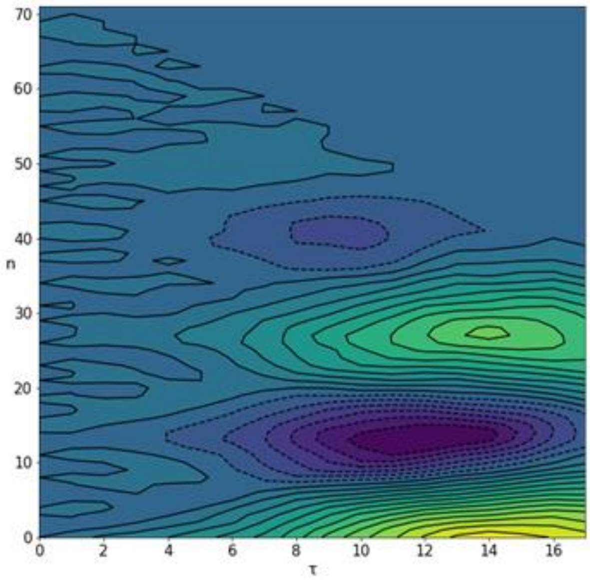

Fig. 10. The autocorrelation function levels map.

The autocorrelation function local maxima arguments along the ordinate axis in Fig. 10 coincide with the local minima obtained from the shear function in Fig. 8. The shear function global minimum and the autocorrelation function global maximum corresponding to 27 years are also distinguished. Thus, it can be expected that the primary energy consumption dynamics in the United States will show temporal rhythms lasting 10 and 27 years.

6 Building a “Development Cell”

The conditions for the rhythms' coordination, built according to the arithmetic progression law, and rhythms following a geometric progression, are formulated in the critical phenomena theory. The corresponding structure was named the development cell [8].

For the time interval from 1635 to 2020, we will construct a development cell along the abscissa axis (see Fig. 11), based on the duration equal to e·34.5≈93.8.

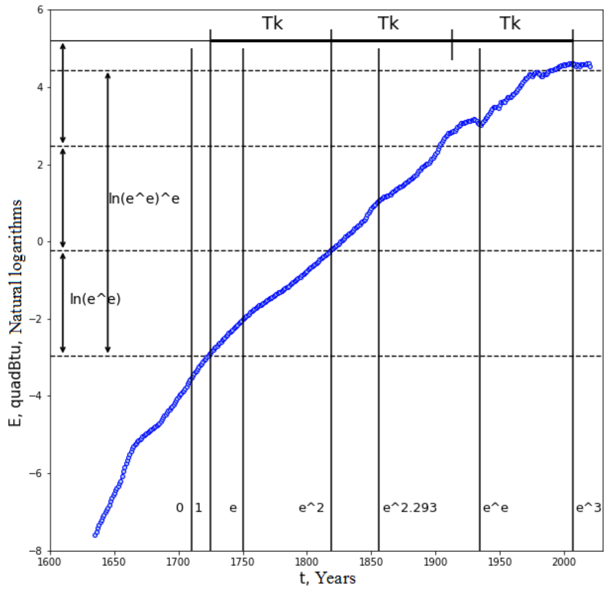

Fig. 11. Rhythmic structure manifested in the energy consumption dynamics.

The development cell, which reconciles the arithmetic and geometric progression in the presented data, is plotted along the ordinate in Fig. 11. At the moment, the energy consumption dynamics in the United States is in the "restructuring phase".

7 ARIMA Autoregressive Model

To obtain predictive estimates, we will use the Box-Jenkins methodology, or the ARIMA autoregressive model [9], which is designed to analyse non-stationary time series. This model takes into account 3 parameters: the autoregression order p, which shows the current period value dependence on past values, the differences' order in the original series d, the moving average q order which allows setting the model error as the previously observed error values a linear combination. These parameters are expressed as whole numbers and take into account seasonality, trend, and noise in the datasets.

The parameters P, D, Q are used to track seasonality. They follow the same definitions as the non-seasonal parameters (p, d, q). The parameter s is also used, which determines the frequency of the time series (4 - quarterly periods, 12 - annual periods). To track seasonality, a modification of the basic model ARIMA (p, d, q) - SARIMA (P, D, Q) is used. The parameters P, D, Q are used to track seasonality. They follow the same definitions as the non-seasonal parameters (p, d, q). The parameter s is also used, which determines the time series frequency (4 - quarterly periods, 12 - annual periods).

The Python programming language library auto.arima () function takes a non-stationary time series and selects the required parameters optimal values. For the energy consumption dynamics time series, the following parameters were obtained: p = 1, d = 1, q = 1, P = 4, D = 1, Q = 1, s = 12.

The applying the autoregressive model result to the data on the primary resources energy consumption general dynamics in the United States is shown in Fig. 12.

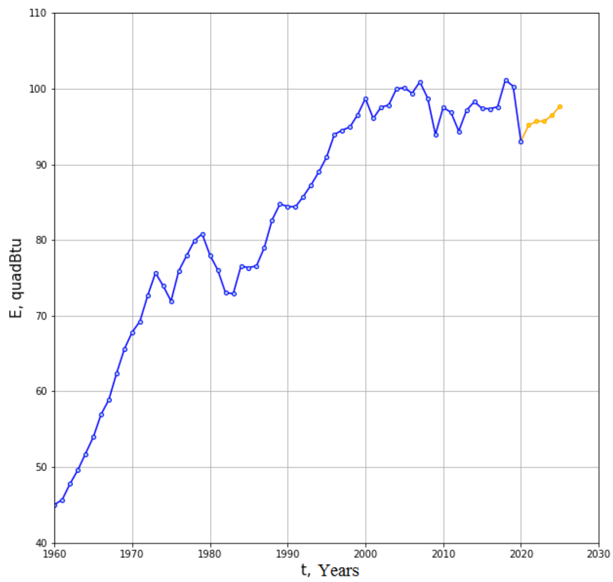

Fig. 12. Forecast based on the ARIMA model.

The resulting forecast estimates based on the ARIMA model for 2021-2025 are presented in Table 3.

Table 3. The energy consumption dynamics forecast estimates in the United States.

Year	Power consumption level, quads Btu
2021	95.17
2022	95.70
2023	95.69
2024	96.49
2025	97.66

8 Conclusion

The long-term trends' analysis in the annual energy consumption dynamics in the United States leads to the following conclusions:

- in the United States of America in the period 1720-1978, with 258 years a total duration, according to an exponential development regime, the primary energy resources annual consumption increased with a constant average growth rate of 2.9% per year, which ensured the American economy main development;

- annual energy consumption in the US has now reached a plateau and is in the high point vicinity. The primary energy resources consumption in the United States is in the midst of an energy crisis and a shortage of resources, which way out is seen in the new technologies' creation for obtaining energy. In 2013, the accumulated energy consumption curve passed the inflexion point, therefore, it was not possible to stabilize the annual energy consumption at the maximum level;
- by 2025, it is planned to locally increase the annual energy consumption level to 97.66 quads Btu;
- the energy consumption share derived from oil and coal is falling, while the energy share from gas and nuclear sources is increasing.

In the study course, a program was developed that passed the Russian Federation software state registration.

References:

[1] Veniaminov S. S., *Revealing Latent Structure Laws in Processes and Signals*, URSS, Moscow, 2014.
[2] Kuzmin V. I., Gadzaov A. F., *Method of Construction of Models on Empirical Data*, MGTU MIREA, Moscow, 2012.
[3] Serebrennikov M. G., Pervozvansky A. A., *Revealing of the Latent Periodicity*, Science, Moscow, 1965.
[4] Dzerjinsky R. I., Pronina E. N., Dzerzhinskaya M. R., The Structural Analysis of the World Gold Prices Dynamics, *Advances in Intelligent Systems and Computing*, Vol.1225, 2020, pp. 352–365.
[5] Paramonov A. A., Kuzmin V. I., Dzerjinsky R. I., Analysis of almost-periodic and almost-proportional characteristics of a representative sample local minima time series, *IOP Conference Series: Materials Science and Engineering*, Vol. 1047, 2021, 012045.
[6] Energy Information Administration, https://www.eia.gov/todayinenergy/detail.php?id=44277.
[7] Melamed I. I., Prokopyeva M. S., Pronina E. N., Energy Strategies of the Asia-Pacific Region: Problems and Prospects (Part 2). *Moscow University Bulletin*, Vol. 25, 2013, pp. 130-168.
[8] Kuzmin V. I., Galusha N. A., Popov S. A., *The Modern Civilization Crisis*, RIOR, Moscow, 2011.
[9] Machine Learning, http://www.machinelearning.ru/wiki/.