



Publisher

<http://jssidoi.org/esc/home>



IMPROVING THE SOCIO-ECONOMIC WELL-BEING OF THE POPULATION IN RURAL AREAS IN THE CONTEXT OF IMPROVING THE ECONOMIC SECURITY OF THE STATE

Oleg V. Zakharchenko

*Department of Management and Marketing, Odessa State Academy of Civil Engineering and Architecture, 65029,
4 Didrikhson Str., Odesa, Ukraine*

E-mail: o.zakharchenko5206@nuos.pro

Received 16 March 2020; accepted 31 August 2020; published 30 December 2020

Abstract. Issues of improving the well-being of the population do not always belong to the category of strategic. In this case, the main value is determined by individual indicators that reflect the development of individual territories or industries. The relevance of the paper is determined by the ability of calculating indicators that form an idea of how well developed the socio-economic program of the territory development. The novelty of the study is determined by the fact that has been developed a model that allows to accumulate a number of formal indicators and thus not only specify the level of development of a particular territory or region, but also indicate the level of participation of private business in the development of socio-economic well-being. The authors include in the model the structure of socio-economic indicators that have a personal basis and can be used to develop programs and systems for computer modelling of the socio-economic development of the territory. The authors show that socio-economic well-being subsequently determines the level of economic security of the population. The practical significance of the study is determined by the capabilities of forecasting and applying indirect models for determining the socio-economic well-being of the population in the conditions of crisis conditions of the economy. It is proposed to use a set of balanced indicators of socio-economic development with a predominance of mathematical modelling patterns based on a set of market data.

Keywords: strategy; model; development; population; well-being

Reference to this paper should be made as follows: Zakharchenko, O.V. 2020. Improving the socio-economic well-being of the population in rural areas in the context of improving the economic security of the state. *Entrepreneurship and Sustainability Issues*, 8(2), 512-532. [http://doi.org/10.9770/jesi.2020.8.2\(31\)](http://doi.org/10.9770/jesi.2020.8.2(31))

JEL Classifications: B55, O15, P10, P16

Make your research more visible, join the Twitter account of ENTREPRENEURSHIP AND SUSTAINABILITY ISSUES:
@Entrepr69728810

1. Introduction

The financial and economic sphere and related processes accompany us at every step of our daily lives (Kline et al., 2007). That is why their study, and especially the study of nonlinear non-stationary processes, requires more and more attention from the scientific community every day (Bryden et al., 2011; Fedotov et al., 2018). In particular, the issues of modelling and forecasting processes that, in the conditions of crisis phenomena of the economy and various external or domestic factors of influence, acquire new features and becoming increasingly relevant (Saqalli et al., 2011; Yessilov, 2015; Salimyanova et al., 2019). To obtain valid mathematical and statistical models, as well as high-quality forecasts, it is necessary to critically analyse, improve and use existing methods, for example, methods based on moving average autoregression (MAA) models, methods of optimal estimation and filtering, the method of group accounting of arguments and many others (Newburn & Berck, 2006). However, in order to achieve a significant improvement of the model and obtain a better forecast, neural networks of various modifications, non-standard regression approaches to modelling can be applied (Rushton, 1992; Bliznets et al., 2018; Pyagay et al., 2014; Bashynska et al., 2019).

The problems of mathematical modelling are becoming increasingly relevant for the effective organisation of management of individual business entities and economic communities due to the fact that the quality of decisions made is largely determined by the quality of forecasting their consequences. Therefore, the decisions made today should be based on reliable estimates of the possible development of the studied phenomena and events in the future (Johnson et al., 2010; Ermilova, 2018; Bashynska et al., 2020).

Forecasting is a tool that represents an opportunity within which realistic objectives for planning a subject's development strategy can be set (Hatherley et al., 2013). During forecasting, mathematical methods are used, which are sometimes quite complex and non-trivial. Due to the widespread use of computers that are easy to understand and use, and most importantly, effective forecasting programs, it has become a fast and fairly effective analysis mechanism (Do & Park, 2018). That is why the search for new or a combination of existing methods allows to get a more accurate and valid model, which in turn provides the opportunity to make effective management decisions (Leinbach & Cromley, 1989).

Today, there are many systems of various nature: environmental, technical, economic and financial, which function and develop under the influence of random disturbances of a different nature. And this in turn leads to the emergence of non-stationary processes – the emergence of a trend and heteroscedasticity (Eid & Hegazy, 1983). Non-stationary, as a rule, are accompanied by non-linear (Moldodikova et al., 1989). This necessitates the development and use of atypical approaches to modeling and forecasting the analysed processes (Gorbenkova et al., 2018). Vivid and well-known examples of such models are socio-economic indicators characterising the economic situation in a country (Refsgaard & Johnson, 2010). They describe non-linear behavior (Masduqi et al., 2010). Here the signs of unsteadiness are (Gronau et al., 2017):

- structure of macroeconomic processes, which is unstable in time;
- uncertainty of the future and critical values of the parameters of control processes;
- complex formalisation of process dynamics.

Provided that not all the features will be taken into account, insufficient statistical sampling is possible when it is either impossible to conduct a correct analysis, assessment and valid forecasting of the dynamics of macroeconomic processes by classical methods, or the result will be of very poor quality (Rafindadi & Kondo, 2018; Kerimov et al., 2018a; Kerimov et al., 2018b).

Make your research more visible, join the Twitter account of ENTREPRENEURSHIP AND SUSTAINABILITY ISSUES:
@Entrepr69728810

2. Materials and Methods

The adequacy criteria of the model make it possible to separately assess the significance of the coefficients of the mathematical model in a statistical sense, to determine the integral error of the model relative to the initial time series, to establish the correlation between the values of the model error (they should not be correlated), and also to determine the degree of adequacy of the model to the physical process as a whole (Križanović et al., 2016). Consider the following statistical parameters: t – Student's statistic. The significance of each regression coefficient in a statistical sense is determined using t -test, which, as a rule, is calculated by all statistical software packages according to the formula:

$$t_a = \frac{\hat{a} - a_o}{SE_a} \quad (1)$$

where \hat{a} – coefficient estimate obtained using the packet; a_o – null hypothesis regarding the value of this coefficient (usually $a_o=0$); SE_a – standard error of the estimate of the coefficient calculated in the packet.

Obviously, the lower the standard error, the better the estimate of the model coefficient (Müller & Munroe, 2005). To determine the significance of the coefficient, it is necessary to know the sample length N , the number of estimated parameters p and set the significance level α (usually set $\alpha=1\%$, $\alpha=5\%$, $\alpha=10\%$). A significance level of 5% means that when assessing the regression, we assume that erroneous decisions about the significance of the estimates are possible in 5% of cases. These parameters allow to select significances from the tables t_{crit} . When $t_{crit} < t_a < t_{crit}$ then the null hypothesis about the insignificance of the coefficient is accepted; otherwise, it is rejected and the coefficient is considered significant. Because the significance of statistics t_a inversely proportional to standard error SE_a , then the greater the significance t_a , the higher the value of a particular coefficient.

Determine coefficient R^2 . The degree of informativeness of the time series often use its variance. Coefficient R^2 – this ratio of the variance of a part of the time series of the main variable is described by the resulting equation to the sample variance of this variable. It is calculated by the formula:

$$R^2 = \frac{var(\hat{y})}{var(y)} \quad (2)$$

For a valid model, the coefficient of determination should tend to unity, that is: $R^2 \rightarrow 1$. The sum of the squared errors of the model $\sum e^2(k)$, that is:

$$SSE = \sum_{k=1}^N [\hat{y}(k) - y(k)]^2 \quad (3)$$

Where N – sample length, measurement (4):

$$\hat{y}(k) = \hat{a}_0 + \hat{a}_1 \hat{y}(k-1) + \hat{a}_2 \hat{y}(k-2) + \hat{b}_1 x(k) + \hat{b}_2 z(k); y(k) \quad (4)$$

Make your research more visible, join the Twitter account of ENTREPRENEURSHIP AND SUSTAINABILITY ISSUES: @Entrepr69728810

Obviously, from the possible candidates it is necessary to choose the model for which $\sum e^2(k)$ takes the minimum value. Akaike Information Criterion (AIC). This criterion takes into account the sum of the squared errors, the number of measurements N and the number of estimated parameters p :

$$AIC = N \ln[\sum_{k=1}^N -e^2(k)] + 2p \quad (5)$$

Obviously, for a better model, the criterion has lesser value, since it depends on the sum of the squared errors (SSE). However, in addition to SSE, this criterion takes into account the sample length and the number of estimated parameters, which makes it more informative. Durbin-Watson statistic. Durbin-Watson statistic are calculated using the formula:

$$DW = 2 - 2\rho \quad (6)$$

where ρ – correlation coefficient between random variable values $\varepsilon(k) \approx e(k)$, that is:

$$\rho = cov[e(k)] = E[e(k)e(k-1)] \quad (7)$$

This parameter allows to determine the correlation degree of model errors. In the complete absence of correlation between the errors $DW = 2$, this is the most acceptable value of this parameter. Fisher statistics F , which determines the validity degree of the model as a whole. For a valid model, the condition: $F > F_{crit}$, where F_{crit} determined by the table similar to t -statistics. Thus, a more valid model corresponds to a larger F value. The Tail coefficient is a very important indicator of model accuracy and compatibility:

$$U = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}}{\sqrt{\frac{1}{N} \sum_{i=1}^N (y_i)^2} + \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{y}_i)^2}} \quad (8)$$

When building, its value is between 0 and 1. If $U = 1$, the model cannot be used for forecasting. Predicted, based on the obtained model, real series are not correlated. Otherwise, if $U = 0$, the predicted series coincide with the real series and the model is most preferable.

3. Results and Discussion

This coefficient can be broken down into the sum of the ratio U^M , the ratio of variations U^S , and the covariance ratio U^C :

$$U^M = \frac{(y_i - \hat{y}_i)^2}{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (9)$$

Make your research more visible, join the Twitter account of ENTREPRENEURSHIP AND SUSTAINABILITY ISSUES: @Entrepr69728810

U^M is used to check for systematic deviations for average real and predicted series. Or, in other words, the model constantly overestimates the forecast. The lower the value U^M the better. If $U^M=0$, then in the predicted values there is no bias and the model is qualitative. U^S – the ratio of variations and is defined as:

$$U^M = \frac{(\sigma_{actual} - \sigma_{fitted})^2}{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (10)$$

The ratio of variations is used to verify that the model has enough dynamic properties to absorb variations of real series. For example, a model can provide systematically smaller oscillations than oscillations of real series. Similarly, U^M , lower values of U^S are an indicator of less variations. U^C – the ratio of covariances and is defined as:

$$U^C = \frac{2(1-\rho)(\sigma_{actual} - \sigma_{fitted})^2}{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (11)$$

Covariance ratio U^C measures how correlated the predicted and real series are. Equivalence of U^C to zero indicates that the predicted and real series are perfectly correlated. It should be noted that:

$$U^C + U^S + U^M = 1 \quad (12)$$

To evaluate the model, it is necessary to determine how well the model reproduces real time series. It is always recommended to make a repeated (retrospective) forecast after modelling. The formal criteria for evaluating the forecast are: formal statistics; turning points (inflection points); sensitivity to changes in the source data; sensitivity to changes in coefficients. Root mean square error (RMSE):

$$RMSE = \sqrt{\frac{1}{S} \sum_{i=1}^S (y(k+s) - \hat{y}(k+s, k))^2} \quad (13)$$

Average forecast error (AFE):

$$AFE = \frac{1}{S} \sum_{i=1}^S y(k+s) - \hat{y}(k+s, k) \quad (14)$$

Average error in percent (PAE):

$$PAE = \frac{1}{S} \sum_{i=1}^S \frac{y(k+s) - \hat{y}(k+s, k)}{y(k+s)} \times 100\% \quad (15)$$

Absolute average error in percent (AAE):

$$AAE = \frac{1}{S} \sum_{i=1}^S \frac{|y(k+s) - \hat{y}(k+s, k)|}{|y(k+s)|} \times 100\% \quad (16)$$

Make your research more visible, join the Twitter account of ENTREPRENEURSHIP AND SUSTAINABILITY ISSUES:
@Entrepr69728810

Maximum absolute error (MAE):

$$MAE = \max\{|y(k+1) - \hat{y}(k+1, k)|, \dots, |y(k+s) - \hat{y}(k+s, k)|\} \quad (17)$$

Minimum absolute error (MIAE):

$$MIAE = \min\{|y(k+1) - \hat{y}(k+1, k)|, \dots, |y(k+s) - \hat{y}(k+s, k)|\} \quad (18)$$

Estimating models by inflection points is an important indicator, as some models may have greater accuracy, but there may be cases where they may not work well in predicting changes in trends or cycles. Other models may be less accurate, but may exhibit a higher dynamic character. We can talk about a compromise between accuracy and dynamic properties. Unfortunately, there is no formula test for these properties. However, a visual check of the predicted and real series quickly determines if the model includes inflection points. Another important test of the quality of the model is the analysis of sensitivity to the initial (starting) data. If the model gives results that are generally roughly independent of the initial data, then this model is considered to be of high quality.

To create an adaptive forecasting system, it is necessary to choose a process, analyse its current state, existing models and approaches to its forecasting. An analysis of specialised literature can greatly help in finding an existing model to describe the behaviour of a selected process. For example, we can find mathematical models in the form of systems of equations, laws of distribution of input and output quantities (statistical models), or logical models in the form of sets of rules that characterise the interaction of inputs and outputs. Recently, probabilistic methods and models of various structures, as well as models in the form of fuzzy logic rules, have become more and more widespread and popular. To implement the further stages of creating predictive and control systems, a choice is made of the type and structure of the model (Kuznetsov et al., 2018; Lapidus et al., 2018a; Lapidus et al., 2018b).

For example, consider a model that is created on the basis of theoretical concepts and patterns regarding a specific process. It may require only clarification of its parameters, which can be obtained using statistical data. And in another case, a model that is completely based on statistical studies may require significantly larger amounts of information and time to build it. It is also necessary to inspect other specialised literature, but we need to know the features of the selected methods and the boundaries of their application. The software product obtained this way is the simplest in terms of architecture. That is why the use of this software product will be appropriate, convenient and profitable in organisations that do not have in-depth knowledge of computer programming. The architecture of the simulated software product has the following levels (Fig. 1).

Make your research more visible, join the Twitter account of ENTREPRENEURSHIP AND SUSTAINABILITY ISSUES:
@Entrepr69728810

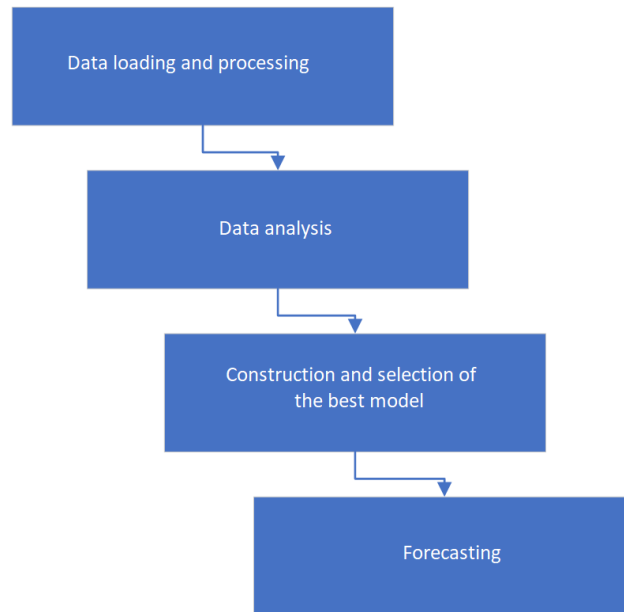


Fig. 1. Architecture levels of the simulated software product

Consider each of these levels: data is loaded at the first level; at the second level, it is possible to visually evaluate data, conduct statistical, correlation analysis, analysis for non-linearity and non-stationarity; at the third level, models are built; at the fourth level, forecasting is based on the created model. It is also possible to depict a functional diagram that expresses the content of the simulated software product (Fig. 2).

To analyse and build models and forecasts, four time series were chosen: USD/CHF (208 values), CPI (60 values), LTV (107 values) and RST (100 values). USD/CHF (US Dollar/Swiss Franc) is a currency pair consisting of the US dollar and Swiss franc. On traders' slang, this pair is also known as Swiss. This pair is in fifth place among the most traded currencies on the market and is the world's reserve currency. USD/CHF is considered a safe haven for traders due to the economic stability and the neutral political nature of Switzerland. As data we take the minimum weekly prices of the USD/CHF currency pair (208 values) from 2016 (Fig. 3).

Make your research more visible, join the Twitter account of ENTREPRENEURSHIP AND SUSTAINABILITY ISSUES:
@Entrepr69728810

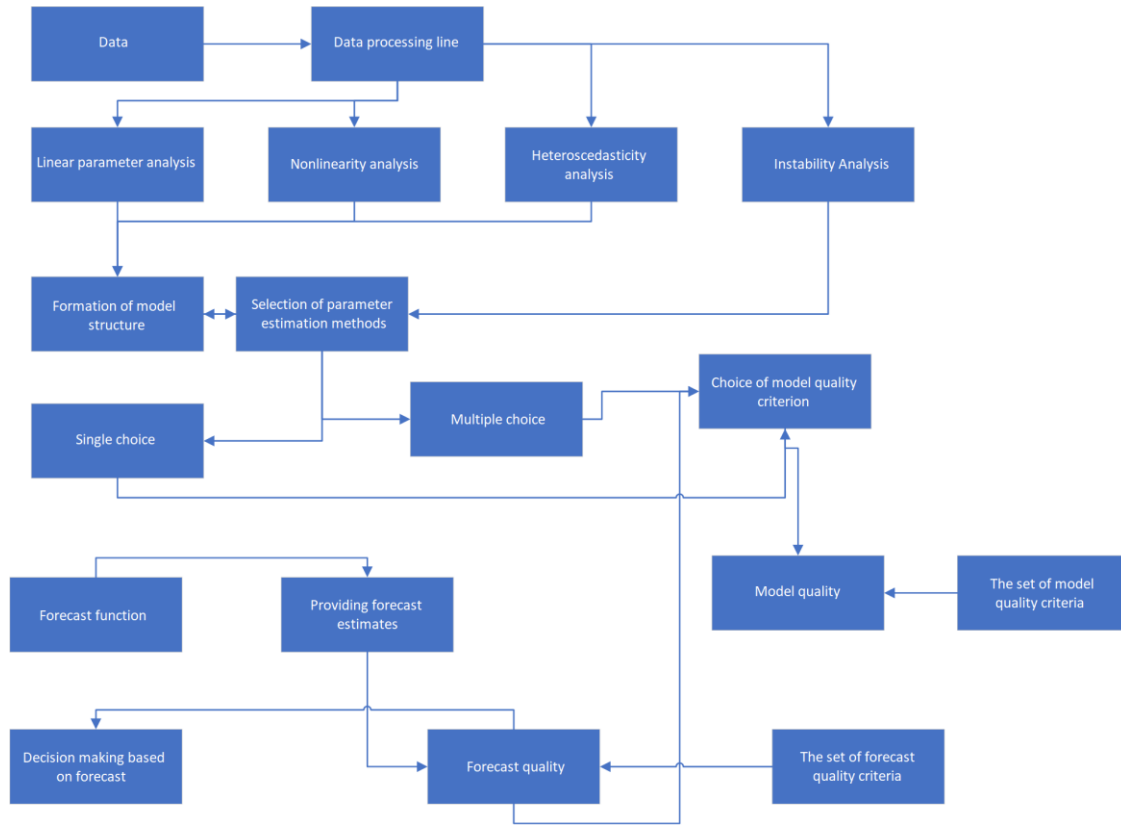


Fig. 2. Functional diagram of a simulated software product

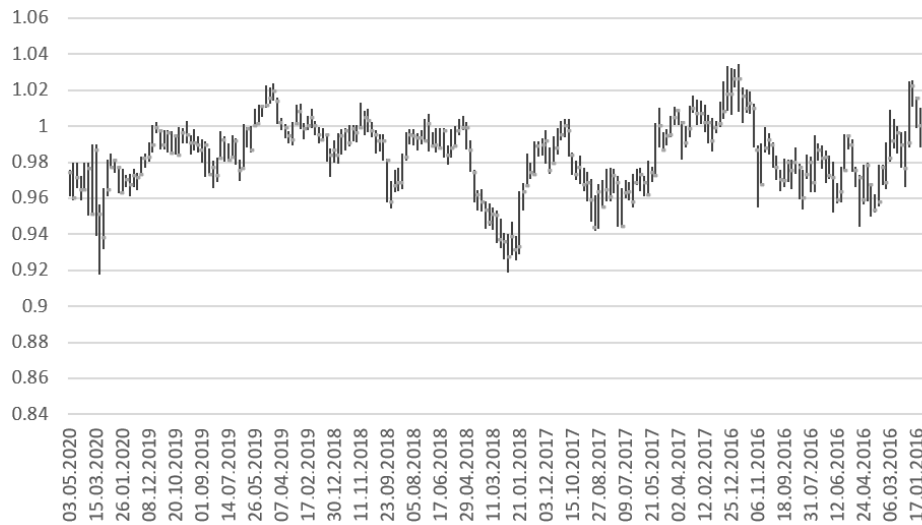


Fig. 3. USD/CHF series

Make your research more visible, join the Twitter account of ENTREPRENEURSHIP AND SUSTAINABILITY ISSUES: @Entrepr69728810

Statistical characteristics of the USD CHF series: average 1.2107; dispersion 0.0005; asymmetry coefficient 0.2803; excess 2.4929; Jacques-Berah 4.9518. CPI – Consumer Price Index (CPI) characterises the changes over time of the general price level for goods and services that the population buys for non-productive consumption. It is an indicator of changes in the cost of a fixed set of consumer goods and services in the current period to its value in the base period.

CPI (Fig. 4) is the most important indicator that characterises inflationary processes in the country's economy and is used to solve many issues of state policy, analyse and forecast price processes in the economy, review the size of cash incomes and minimum social guarantees of the population, resolve legal disputes, recount indicators system of national accounts at constant prices.

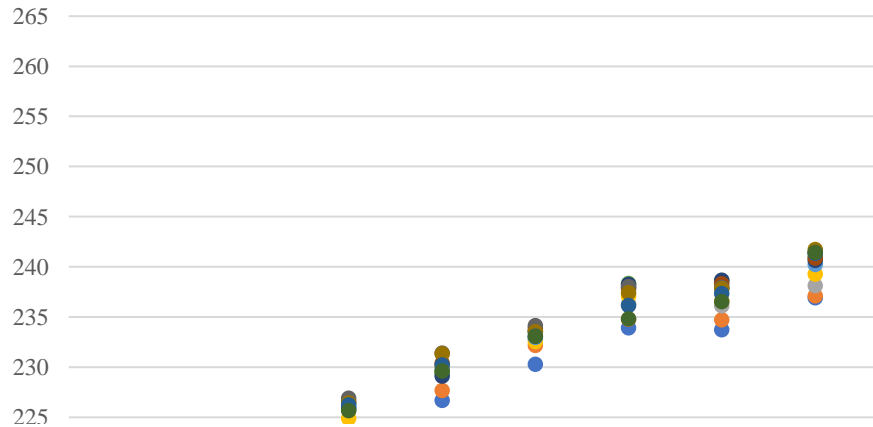


Fig. 4. CPI series (consumer price index)

Statistical characteristics of the CPI series: average 100.35; dispersion 1.5534; asymmetry coefficient 0.2599; excess 3.4014; Jacques-Berah 1.0785. Gross revenue is the amount of money a company receives as a result of doing business, and expenses are not taken into account. For example, a retailer receives revenue from the sale of goods, but the gross revenue does not include the cost of acquiring goods from suppliers, labor costs for employees, overhead associated with running a business, and the lost revenue from returning goods or theft. Gross revenue refers only to the amount of income received only from the sale of goods or services, or from other sources of income, such as royalties and investments, excluding related costs. It should be noted that gross revenue is not the main indicator for many companies, although in some situations it is of great importance (Fig. 5).

Make your research more visible, join the Twitter account of ENTREPRENEURSHIP AND SUSTAINABILITY ISSUES:
@Entrepr69728810

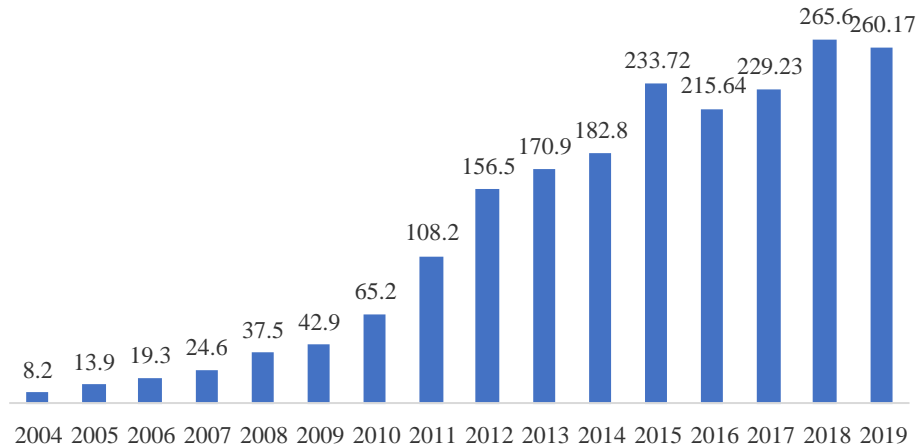


Fig. 5. GR series (Gross Revenue) Apple Inc, 2004-2019

Since this indicator does not reflect the costs associated with doing business, it does not always reflect how efficiently the company operates. In professional language, gross revenue is sometimes called the “Top Line”, since this indicator is usually placed at the top of various financial statements. At the same time, expense items associated with the conduct of activities are placed below this indicator (Fig. 6). Statistical characteristics of Gross Revenue: average 14380.36; dispersion 527119888.4945; asymmetry coefficient 0.7334; excess 2.7167; Jacques-Berah 10.0439. RST – net profit.

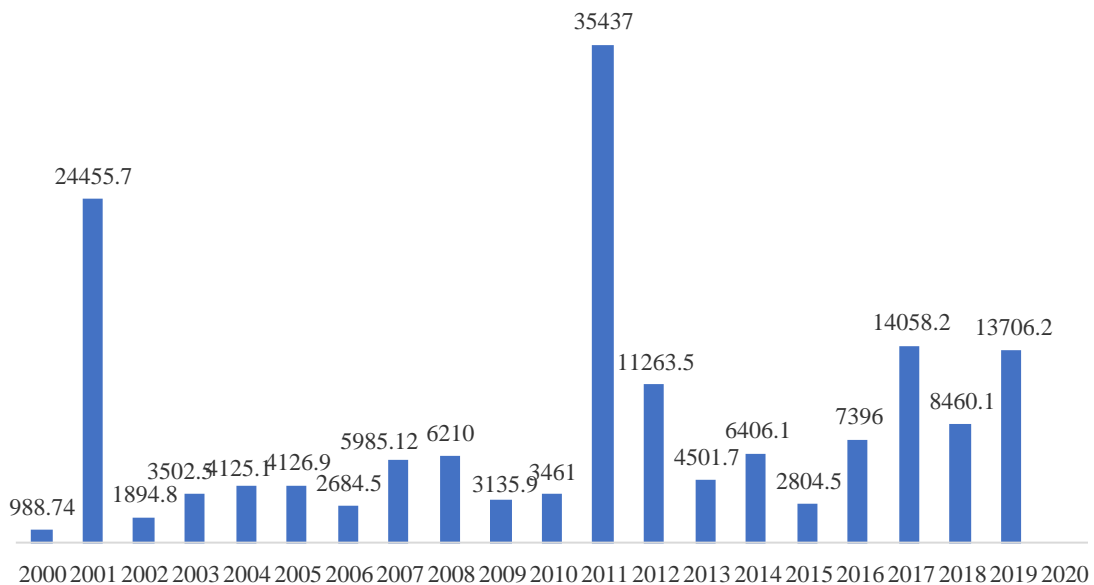


Fig. 6. RST series (average per employee, 2000-2020)

Make your research more visible, join the Twitter account of ENTREPRENEURSHIP AND SUSTAINABILITY ISSUES: @Entrepr69728810

Statistical characteristics of the RST series: average 0.2634; dispersion 23.69521; asymmetry coefficient 0.5483; excess 4.3794; Jacques-Berah 12.9377. First, consider the possibility of describing the selected processes using the autoregressive model. Autoregressive models are the simplest in their structure, but quite often they have a high degree of validity to the process under study, acceptable for further use. We build an autoregressive model for the USD/CHF series: autoregression of the 14-th order:

$$\begin{aligned} y(k) = & 0,0133 + 0,9655*y(k-1) + 0,0099*y(k-2) + \\ & + 0,0319*y(k-3) - 0,1146*y(k-4) + 0,0521*y(k-5) + \\ & + 0,1001*y(k-6) - 0,0682*y(k-7) - 0,0918*y(k-8) + \\ & + 0,0995*y(k-9) + 0,1130*y(k-10) - 0,1521*y(k-11) + \\ & + 0,0017*y(k-12) + 0,0472*y(k-13) - 0,0057*y(k-14) \end{aligned} \quad (19)$$

Therefore, statistical characteristics of the model: $R^2 = 0,9516$; $\sum e^2 = 0,00459$; $DW = 1,9692$. The determination coefficient takes on values (0.9516), the sum of the squared errors is very small (0.00459), and the Durbin-Watson statistic (1.9692) is approaching the best value. The quality characteristics (for the training sample) for the forecast of the selected series RMSE = 0.003473; PAE = 4.2%. (Fig. 7):

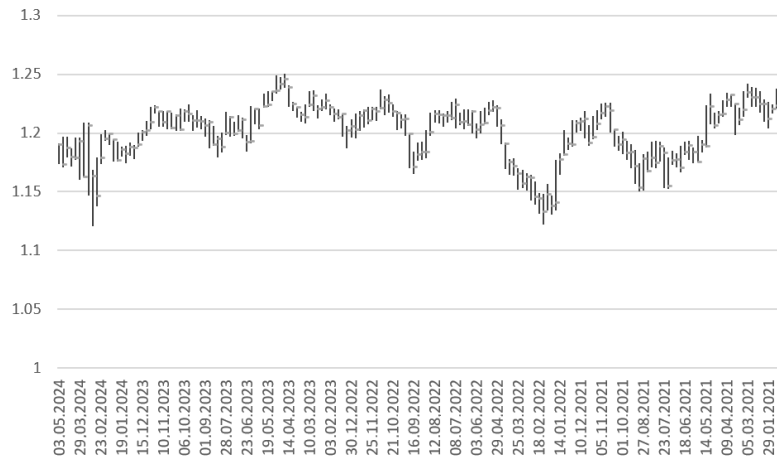


Fig. 7. Forecast for USD/CHF

That is, the root means square error (RMSE) and the average error in percent (PAE) indicate the general suitability of the forecasting model. Next, we will build another model for the USD/CHF series: 5-th order autoregression:

$$\begin{aligned} y(k) = & 0,0140 + 0,9637 * y(k-1) - 0,0114 * y(k-2) + 0,0832 * \\ & * y(k-3) - 0,1271 * y(k-4) + 0,0796 * y(k-5) \end{aligned} \quad (20)$$

Make your research more visible, join the Twitter account of ENTREPRENEURSHIP AND SUSTAINABILITY ISSUES: @Entrepr69728810

Therefore, statistical characteristics of the model: $R^2 = 0,9541$; $\sum e^2 = 0,00463$; $DW = 2,0024$. The coefficient of determination takes on values (0,9541), the sum of the squared errors is very small (0.00463), and the Durbin-Watson statistic (2.0024) is approaching the best value. The quality characteristics (for the selected row RMSE=0,003778; PAE=4,524% (Fig. 8).

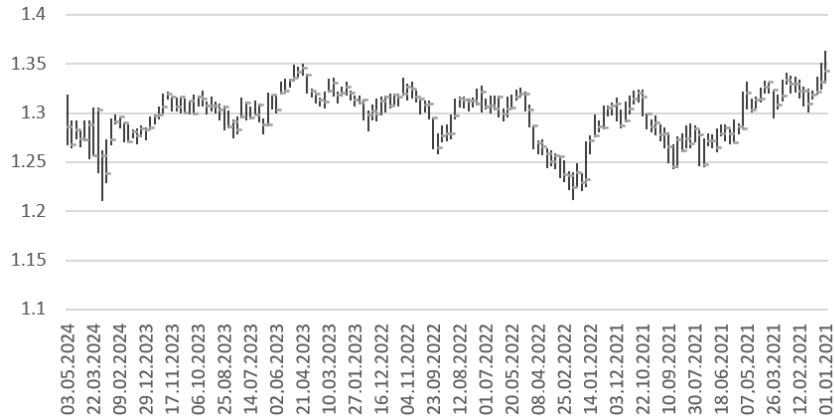


Fig. 8. Forecast for USD/CHF

That is, the root means square error (RMSE) and the average error in percent (PAE) indicate the general suitability of the model for forecasting. Here is a general table of model quality estimates and forecast for a series of USD/CHF (Table. 1).

Table 1. Results of modelling and forecasting for a series of USD/CHF according to various quality criteria

Statistical series	Model type	Model validity			Forecast characteristic	
		R^2	$\sum e^2$	DW	RMSE	PAE %
USD/CHF	AR(14)	0.9516	0.00459	1.9692	0.003473	4.2
	AR(5)	0.9541	0.00463	2.0024	0.00378	4.524

As can be seen from the table, the AR showed the best in terms of forecast quality and model quality (14). We will build an autoregressive model for the CPI series: 15-th order autoregression:

$$\begin{aligned}
 y(k) = & 43,266 + 0,6589*y(k-1) - 0,4392*y(k-2) + 0,4476*y(k-3) - \\
 & -0,1289*y(k-4) - 0,0756*y(k-5) + 0,0068*y(k-6) - 0,0273*y(k-7) - \\
 & -0,0604*y(k-8) - 0,0178*y(k-9) + 0,1186*y(k-10) - 0,0039*y(k-11) + \\
 & +0,5238*y(k-12) - 0,3695*y(k-13) + 0,1725*y(k-14) - \\
 & -0,2428*y(k-15)
 \end{aligned}
 \tag{21}$$

Therefore, statistical characteristics of the model: $R^2 = 0,544$; $\sum e^2 = 0,696$; $DW = 2,035$. The determination coefficient takes on values (0.544), the sum of the squared errors is very small (0.696), and the

Make your research more visible, join the Twitter account of ENTREPRENEURSHIP AND SUSTAINABILITY ISSUES: @Entrepr69728810

Durbin-Watson statistic (2.035) is approaching the best value. The quality characteristics (for the training selection) for the forecast of the selected series RMSE = 1,0504; PAE = 3.28% (Fig. 9).

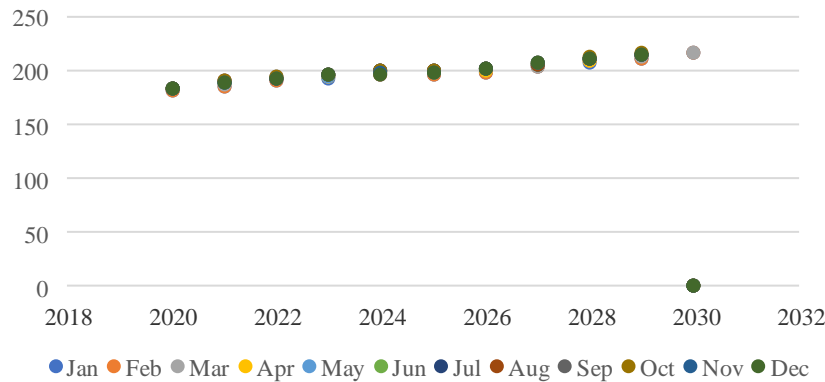


Fig. 9. Forecast for CPI

That is, the root means square error (RMSE), the average error in percent (PAE) indicates the general suitability of the model for forecasting. Next, it is necessary to build another autoregressive model for the CPI series: autoregression of 3-rd order with a moving average value of 5:

$$y(k) = 27,1206 + 1,1446*y(k-1) - 0,7030*y(k-2) + 0,2875*y(k-3) - 0,4281*v(k-1) + 0,0572*v(k-2) + 0,2875*v(k-3) - 0,156*v(k-4) - 0,3263*v(k-5) \quad (22)$$

Therefore, statistical characteristics of the model: $R^2 = 0,9496$; $\sum e^2 = 1,038$; $DW = 2,7359$. The determination coefficient takes on values (0.9496) better than the previous one, the sum of the squared errors (1.038) shows a certain deterioration according to preliminary results, and the Durbin-Watson statistic (2.7359) showed poor results and a significant deterioration with previous models. The quality characteristics (for the training selection) for the forecast of the selected series RMSE = 0.909; PAE = 2.81% (Fig. 10).

Make your research more visible, join the Twitter account of ENTREPRENEURSHIP AND SUSTAINABILITY ISSUES:
@Entrepr69728810

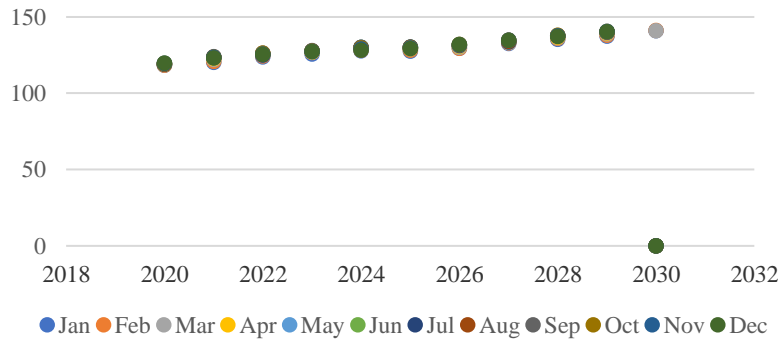


Fig. 10. Forecast for CPI

The root means square error (RMSE), the average error in percent (PAE) decreased compared to the previous result. Here is a general table of model quality estimates and forecast for a CPI series (Table. 2):

Table 2. Results of modelling and forecasting of the consumer price index by various quality criteria

Statistical series	Model type	Model validity			Forecast characteristic	
		R^2	$\sum e^2$	DW	RMSE	PAE%
CPI	AR(14)	0.544	0.696	2.035	1.0504	3.28
	MAA(3.5)	0.9496	1.038	2.7359	0.909	2.81

The constructed MAA model (3.5) had good results, but the Durbin-Watson statistic was 2.7359, which is a big deviation, but the forecast is of high-quality. We will build an autoregressive model for the RST series: 14-th order autoregression:

$$\begin{aligned}
 y(k) = & 0,0974 + 1,1098y(k-1) + 0,1776*y(k-2) + \\
 & +0,0297*y(k-3) - 0,0054*y(k-4) + 0,1289*y(k-5) + \\
 & +0,0809*y(k-6) - 0,0747*y(k-7) - 0,1301*y(k-8) + \\
 & +0,3111y(k-9) + 0,0717y(k-10) - 0,0392*y(k-11) + \\
 & +0,1999*y(k-12) + 0,3538*y(k-13) - 0,367*y(k-14)
 \end{aligned} \quad (23)$$

Therefore, statistical characteristics of the model: $R^2 = 0,91$; $\sum e^2 = 1,567$; $DW = 1,946$. The determination coefficient takes on values (0.91), the sum of the squared errors (1.567), and the Durbin-Watson statistic (1.946). The quality characteristics (for the training selection) for the forecast of the selected series RMSE=2.22; PAE=3.84% (Fig. 11):

Make your research more visible, join the Twitter account of ENTREPRENEURSHIP AND SUSTAINABILITY ISSUES:
@Entrepr69728810

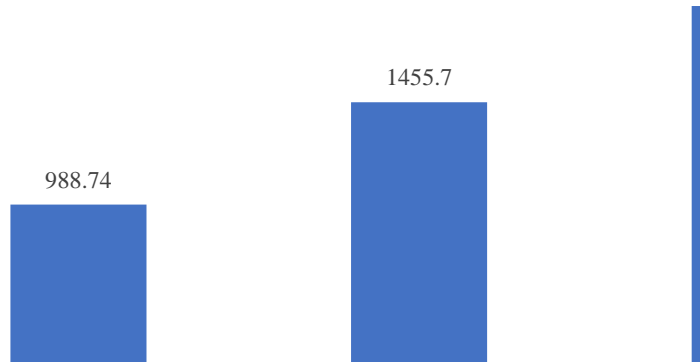


Fig. 11. Forecast for RST

Next, we will build another autoregressive model for the RST series: autoregression of the 5-th order:

$$y(k) = 0,0504 + 1,0972 * y(k - 1) - 0,1664 * y(k - 2) + 0,0086 * y(k - 3) - 0,0031 * y(k - 4) + 0,118 * y(k - 5) \quad (24)$$

Therefore, statistical characteristics of the model: $R^2 = 0,881$; $\sum e^2 = 1,7161$; $DW = 2,0003$. The coefficient of determination takes on the value (0.881), a certain deterioration with the previous results, the sum of the squared errors takes on the value (1.7161), and the Durbin-Watson statistic (2,0003). The quality characteristics (for the training selection) for the forecast of the selected series RMSE=3.084; PAE=5.7% (Fig. 12):

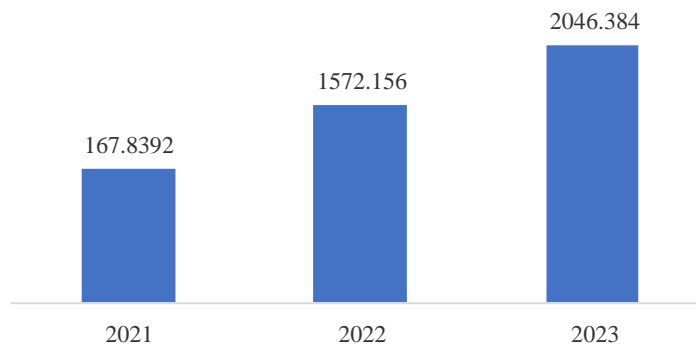


Fig. 12. Forecast for RST

Next, build another autoregressive model for the RST series: autoregression of the 14-th order with a moving order average of 6:

Make your research more visible, join the Twitter account of ENTREPRENEURSHIP AND SUSTAINABILITY ISSUES: @Entrepr69728810

$$y(k) = -1,694 + 0,7666*y(k-1) - 0,2212*y(k-2) + 0,0918*y(k-3) + 0,6318*y(k-4) - 0,0669*y(k-5) - 0,0281*y(k-6) - 0,3432*y(k-7) - 0,0723*y(k-8) + 0,1748*y(k-9) - 0,0125*y(k-10) + 0,1072*y(k-11) - 0,0728*y(k-12) + 0,2445*y(k-13) - 0,447*y(k-14) + 0,3569*v(k-1) + 0,3941*v(k-2) + 0,3301*v(k-3) - 0,4632*v(k-4) - 0,3127*v(k-5) - 0,3943*v(k-6) \quad (25)$$

Therefore, statistical characteristics of the model: $R^2 = 0,9357$; $\sum e^2 = 2,0124$; $DW = 1,144$. The determination coefficient takes on values (0.9357), the sum of squared errors (2.0124), and Durbin-Watson statistic (1.144) showed a significant deterioration in the results. The quality characteristics (for the training selection) for the forecast of the selected series: RMSE=4.306; PAE=6.18%. The root means square error (RMSE), the average error in percent (PAE) decreased compared to the previous result. Since the series is heteroscedastic, we construct a 5th-order autoregressive model with a quadratic trend (Table 3).

Table 3. Evaluation results of the AR (5) model with a quadratic trend for the RST series

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	0.540648	0.742731	0.727919	0.4686
C(2)	1.060074	0.107005	9.906769	0.0000
C(3)	-0.168990	0.156169	-1.082099	0.2822
C(4)	0.009839	0.158695	0.061999	0.9507
C(5)	0.005448	0.157561	0.034579	0.9725
0(6)	-0.025252	0.107743	-0.234377	0.8152
C(7)	-0.040421	0.034520	-1.170945	0.2448
0(8)	0.000466	0.000329	1.415894	0.1604
R-squared	0.886090	Mean dependent var		0.236716
Adjusted R-squared	0.876925	S.D. dependent var		5.019531
S.E. of regression	1.760954	Akaike info criterion		4.050040
Sum squared resid	29.78344	Schwarz criterion		4.265104
Log likelihood	-184.3769	Hannan-Quinn criterion		4.136942
F-statistic	96.68033	Durbin-Watson statistic		2.001821
Prob(F-statistic)	0.000000			

As a result, the following model was obtained:

$$y(k) = 0,541 + 1,06*y(k-1) - 0,169*y(k-2) + 0,0098*y(k-3) + 0,0054*y(k-4) - 0,02525*y(k-5) - 0,040*k + 0,00046*k^2 \quad (26)$$

The validity of the model is estimated by the following parameters: $R^2 = 0,886$; $\sum e^2 = 29,7834$; $DW = 2,001$. The determination coefficient takes on values (0.886), the sum of the squared errors (29.7834) showed a big value, and the Durbin-Watson statistic (2.001) are almost ideal. The

Make your research more visible, join the Twitter account of ENTREPRENEURSHIP AND SUSTAINABILITY ISSUES: @Entrepr69728810

quality characteristics (for the training selection) for the forecast of the selected series RMSE=3.948; PAE=5.9%. Here is a general table of model quality estimates and forecast for a number of RST (Table 4).

Table 4. Results of modelling and forecasting of RST series

Statistical series	Model type	Model validity			Forecast characteristic	
		R^2	$\sum e^2$	DW	RMSE	PAE%
RST	AR(14)	0.91	1.567	1.946	2.22	3.84
	AR(5)	0.881	1.7161	2.0003	3.084	5.7
	MAA(14.6)	0.9357	2.0124	1.144	4.306	6.18
	AR(5)+k2	0.886	29.7834	2.001	3.948	5.9

The best results in forecasting, i.e., with the slightest errors, were shown by the AR (14) model. We build an autoregressive model for the GR series: autoregression of the 13-th order:

$$y(k) = -23,2267 + 0,8006*y(k-1) - 0,0204*y(k-2) + 0,1980*y(k-3) - 0,0778*y(k-4) - 0,1195*y(k-5) - 0,0915*y(k-6) + 0,049*y(k-7) + 0,1996*y(k-8) - 0,019*y(k-9) + 0,0661*y(k-10) - 0,0583*y(k-11) + 0,6971*y(k-12) - 0,5836*y(k-13) \quad (27)$$

Therefore, statistical characteristics of the model: $R^2 = 0,97$; $\sum e^2 = 1180,84$; $DW = 2,09$. The coefficient of determination takes the value (0.97), the sum of the squared errors (1180.84), and the Durbin-Watson statistic (2.09). The quality characteristics (for the training selection) for forecasting the selected series RMSE=1760.38; PAE=1.88%. Next, we build a different autoregressive model for the GR series: autoregression of the 9-th order:

$$y(k) = -58,2995 + 0,6005*y(k-1) + 0,1344*y(k-2) + 0,2679*y(k-3) + 0,0205*y(k-4) - 0,2479*y(k-5) - 0,3625*y(k-6) + 0,1359*y(k-7) + 0,3401*y(k-8) - 0,1689*y(k-9) \quad (28)$$

Therefore, statistical characteristics of the model: $R^2 = 0,954$; $\sum e^2 = 1515,07$; $DW = 2,0013$. The determination coefficient takes the value (0.954), a certain deterioration with previous results, the sum of the squared errors takes the value (1515.07), and the Durbin-Watson statistic (2.0013). The quality characteristics (for the training selection) for the forecast of the selected series RMSE=2947.597; PAE=3.152%. Next, we build another autoregressive model for the GR series: autoregression of the 13-th order with an average order value of 2:

$$y(k) = 101,5948 + 1,1018*y(k-1) + 0,0192*y(k-2) + 0,0438*y(k-3) - 0,1646*y(k-4) - 0,1465*y(k-5) - 0,0244*y(k-6) + 0,1603*y(k-7) + 0,2244*y(k-8) - 0,1407*y(k-9) - 0,0131*y(k-10) - 0,0728*y(k-11) + 0,7389*y(k-12) - 0,7314*y(k-12) - 0,3671*v(k-1) - 0,3632*v(k-2) \quad (29)$$

Make your research more visible, join the Twitter account of ENTREPRENEURSHIP AND SUSTAINABILITY ISSUES: @Entrepr69728810

Therefore, statistical characteristics of the model: $R^2 = 0,987$; $\sum e^2 = 1296,48$; $DW = 2,393$. The determination coefficient takes on values (0.987), the sum of squared errors (1296.48), and Durbin-Watson statistic showed a significant improvement in the results. The quality characteristics for the forecast of the selected series RMSE=3315.991; PAE=8.4%. The root means square error (RMSE), the average error in percent (PAE) remained with a positive result. Here is a general table of model quality assessments for forecasting for GR series (Table 5),

Table 5. Results of modelling and forecasting of GR series

Statistical series	Model type	Model validity			Forecast characteristic	
		R^2	$\sum e^2$	DW	RMSE	PAE%
GR	AR(13)	0.97	1180.84	2.09	1760.38	1.88
	AR(9)	0.954	1515.07	2.0013	2947.59	3.152
	MAA(13.2)	0.987	1296.48	2.393	3315.99	8.4

The best results in the forecasting, that is, the AR (13) model with small errors. Based on the results of previous studies, we build models of selected processes using the commercial product GMDH Shell, which uses the group argument method to build models. We analyse and predict the range of USD_CHF. Obtained model:

$$Y1[t] = -0.44331 + time * (-0.000169013) + "x1[t - 12], cubert" * 1.56472 \quad (30)$$

The coefficient of determination takes on the value of 0.628. The forecast characteristic of this series assumes the following values RMSE=0.498; PAE=1.278%. We analyse and forecast the consumer price index. Obtained model:

$$Y1[t] = 316.384 + "x1[t - 5], cubert" * (-37.8545) + "x1[t - 3], cubert" * (-8.68243) \quad (31)$$

The coefficient of determination takes on the value of 0.45712. The forecast characteristic of this series assumes the following values RMSE=0.7747; PAE=0.61524%. As we can see, the forecast indicators give a high average absolute error in percent (PAE) which amounted to only 0.615%. It can be concluded that the model is suitable for use. Here is a model that is built for the RST series:

$$Y1[t] = 0.998017 + "x1[t - 1], cubert" * 3.22661 + "x1[t - 3], cubert" * 1.09025 + "x1[t - 8], cubert" * (-1.20318) + "x1[t - 11], cubert" * (-0.428972) + "x1[t - 12], cubert" * 0.412343 + "x1[t - 4], cubert" * 0.506062 \quad (32)$$

The determination coefficient of this model is 0.916219. The forecast characteristic of the RST series is as follows: RMSE=2.66394; PAE=5.27816%. Unfortunately, the forecast indicators are not sufficient, the PAE is 5.27816%. Consider the GR series:

$$Y1[t] = -4293.15 + "x1[t - 1]" * x1[x - 12] * 3.99777e - 05 + x1[t - 7] * x1[t - 8], cubert" * 0.0162273 + x1[t - 3] * x1[t - 10] * 1.29137e - 05 + x1[t - 2] * x1[t - 3], cubert" * (-0.00356083) \quad (33)$$

Make your research more visible, join the Twitter account of ENTREPRENEURSHIP AND SUSTAINABILITY ISSUES: @Entrepr69728810

The determination coefficient of this model is 0.975549. The forecast characteristics of the GR series are as follows: RMSE=816.485; PAE=1.5856%. Here is a summary table for the estimates of the selected series (Table 6).

Table 6. Summary estimation table of selected series

Statistical series	Model type	Model validity			Forecast characteristic	
		R^2	$\sum e^2$	DW	RMSE	PAE%
USDCHF	AR(14)	0.9516	0.00459	1.9692	0.003473	4.2
	AR(5)	0.9541	0.00463	2.0024	0.00378	4.524
	GMDH	0.628			0.498	1.278
CPI	AR(15)	0.544	0.696	2.035	1.0504	3.28
	MAA(3,5)	0.9496	1.038	2.7359	0.909	2.81
	GMDH	0.45712			0.7747	0.61524
RST	AR(14)	0.91	1.567	1.946	2.22	3.84
	AR(5)	0.881	1.716	2.0003	3.084	5.7
	MAA(14,6)	0.9357	2.012	1.144	4.306	6018
	AR(5)+k2	0.886	29.783	2.001	3.948	5.9
	GMDH	0.91621			2.6639	5.2781
GR	AR(13)	0.97	1180.84	2.09	1760.38	1.88
	AR(9)	0.954	1515.07	20013	2947.59	3.152
	MAA(13,2)	0.987	1296.48	2.393	3315.99	8.4
	GMDH	0.976			816.485	1.5856

Conclusions

In order to maximise the reliability and accuracy of the forecast values, modelling was carried out on the basis of several models, after which a corresponding comparative analysis was performed. Prediction of time series was carried out on the basis of autoregressive models, and autoregressive models with moving average, and the method of group accounting of arguments. As a result of modelling based on these methods, it can be concluded that there are no universal methods for constructing and predicting time series. Each row is unique in its own way, and each one requires its own balanced approach. It can also be concluded that during modelling it is impossible to rely only on one parameter of model validity, each of these parameters is important for the calculations and forecasting.

The quality of the constructed model largely depends on the correct application of methods for estimating model parameters. As in linear and non-linear processes, parameters can be estimated using the least squares method, the maximum likelihood method, the Monte Carlo method for Markov chains. Nonlinear methods require setting initial conditions and controlling the convergence of the estimation process. The Monte Carlo method for Markov chains is characterised by flexibility of application, but it may take a long time to implement.

References

Bashynska, I.; Kovalova, O.; Malovichko, O.; Shirobokova, O. 2020. Risk management of innovative socially significant projects (on the example of urban passenger transport), *International Journal of Advanced Research in Engineering and Technology*, 11(4): 294-305.

Make your research more visible, join the Twitter account of ENTREPRENEURSHIP AND SUSTAINABILITY ISSUES:
@Entrepr69728810

Bashynska, I.; Sokhatska, O.; Stepanova, T.; Malanchuk, M.; Rybianets, S.; Sobol, O. 2019. Modelling the risks of international trade contracts, *International Journal of Innovative Technology and Exploring Engineering*, 8(11): 2815-2820.

Bliznets, I.A.; Kartschiya, A.A.; Smirnov, M.G. 2018. Technology transfer in digital era: legal environment, *Tarih Kültür ve Sanat Araştırmaları Dergisi – Journal of History Culture and Art Research*, 7(1): 354-363.

Bryden, J.M.; Johnson, T.; Thomson, K.J.; Ferenczi, T. 2011. Modelling multifunctionality, territorial development, and policy scenarios in rural Europe: An alternative perspective on CAP reform debates, *EuroChoices*, 10(1): 9-16. <https://doi.org/10.1111/j.1746-692X.2011.00145.x>

Do, M.H.; Park, S.C. 2018. New rural development and hierarchical governance in Vietnam: Impacts of government support on rural households' income using a hierarchical linear modelling, *Agris On-Line Papers in Economics and Informatics*, 10(4): 3-15. <https://doi.org/10.7160/aol.2018.100401>

Eid, S.M.A.; Hegazy, O.M. 1983. Investigation into communication needs for rural development through modelling techniques, *Resource Management and Optimization*, 2(3): 269-292.

Ermilova, M. 2018. The use of European experience in mortgage lending in Russian conditions, *International Relations 2018: Current Issues of World Economy and Politics*, 1: 167-175. <https://www.revistaespacios.com/a19v40n13/a19v40n13p27.pdf>

Fedotov, A.M.; Murzakhmetov, A.N.; Dyusembaev, A.E. 2018. Expansion of ideas and processes in social and biological communities, *Eurasian Journal of Mathematical and Computer Applications*, 6(4): 17-28.

Gorbenkova, E.; Shcherbina, E.; Belal, A. 2018. Rural areas: Critical drivers for sustainable development, *IFAC-PapersOnLine*, 51(30): 786-790. <https://doi.org/10.1016/j.ifacol.2018.11.195>

Gronau, S.; Winter, E.; Grote, U. 2017. Modelling nature-based tourism impacts on rural development and conservation in Sikunga Conservancy, Namibia, *Development Southern Africa*, 34(3): 276-294. <http://hdl.handle.net/10.1080/0376835X.2016.1269638>

Hatherley, S.; Cole, W.; Counsell, J.; Geens, A.; Littlewood, J.; Sinnett, N. 2013. An exploration of design alternatives using dynamic thermal modelling software of an exemplar, affordable, low carbon residential development constructed by a registered social landlord in a rural area of Wales, *Smart Innovation, Systems and Technologies*, 22: 601-611. https://doi.org/10.1007/978-3-642-36645-1_56

Johnson, T.G.; Roberts, D.; Wojan, T.R. 2010. Model-based evaluation of rural development policies, *EuroChoices*, 9(1): 30-36. [10.1111/j.1746-692X.2010.00152.x](http://dx.doi.org/10.1111/j.1746-692X.2010.00152.x)

Kerimov, V.; Rachinsky, M.; Mustaev, R.; Serikova, U. 2018a. Geothermal conditions of hydrocarbon formation in the South Caspian basin, *Iranian Journal of Earth Sciences*, 10(1): 78-89.

Kerimov, V.Yu.; Mustaev, R.N.; Osipov, A.V. 2018b. Peculiarities of hydrocarbon generation at great depths in the crust, *Doklady Earth Sciences*, 483(1): 1413-1417.

Kline, J.D.; Moses, A.; Lettman, G.J.; Azuma, D.L. 2007. Modeling forest and range land development in rural locations, with examples from eastern Oregon, *Landscape and Urban Planning*, 80(3): 320-332.

Križanović, Č.V.; Žagar, D.; Rimac-Drlje, S. 2016. Broadband ecosystem elements in techno-economic modelling and analysing of broadband access solutions for rural areas, *Tehnicki Vjesnik*, 23(1): 279-290. <https://doi.org/10.17559/TV-20150508131935>

Kuznetsov, N.B.; Kerimov, V.Yu.; Osipov, A.V.; Bondarev, A.V.; Monakova, A.S. 2018. Geodynamics of the Ural foredeep and geomechanical modeling of the origin of hydrocarbon accumulations, *Geotectonics*, 52(3): 297-311.

Lapidus, A.L.; Kerimov, V.Yu.; Mustaev, R.N.; Movsumzade, E.M.; Salikhova, I.M.; Zhagfarov, F.G. 2018a. Natural bitumens: physicochemical properties and production technologies, *Solid Fuel Chemistry*, 52(6): 344-355.

Lapidus, A.L.; Kerimov, V.Yu.; Tret'yakov, V.F.; Talyshinskii, R.M.; Ilolov, A.M.; Movsumzade, E.M. 2018b. Extraction of asphaltite with toluene, *Solid Fuel Chemistry*, 52(4): 256-259.

Make your research more visible, join the Twitter account of ENTREPRENEURSHIP AND SUSTAINABILITY ISSUES:
@Entrepr69728810

Leinbach, T.R.; Cromley, R.G. 1989. Modeling integrated development investments in rural areas: An Indonesian illustration, *International Regional Science Review*, 12(2): 229-243. <https://doi.org/10.1177/016001768901200207>

Masduqi, A.; Endah, N.; Soedjono, E.S.; Hadi, W. 2010. Structural equation modeling for assessing of the sustainability of rural water supply systems, *Water Science and Technology: Water Supply*, 10(5): 815-823. <https://doi.org/10.2166/ws.2010.339>

Moldodikova, N.I.; Petrov, P.V.; Serbenyuk, S.N. 1989. Automated mathematical cartographic modelling of the interconnected development of urban and rural settlements in the Central Economic Region, *Vestnik Moskovskogo Universiteta, Seriya Geografiya*, 3: 71-78.

Müller, D.; Munroe, D.K. 2005. Tradeoffs between rural development policies and forest protection: Spatially explicit modeling in the Central Highlands of Vietnam, *Land Economics*, 81(3): 412-425.

Newburn, D.A.; Berck, P. 2006. Modeling suburban and rural-residential development beyond the urban fringe, *Land Economics*, 82(4): 481-499. <https://doi.org/10.3368/le.82.4.481>

Pyagay, A.A.; Tulegenova, Z.U.; Urazgaliyeva, M.A.; Zhunusova, Z.G.; Shintayeva, S.S.; Aytzhanova, A. 2014. Complex analysis of risks and statistic dependents in domestic bank system, *Life Science Journal*, 11(6S): 133-139.

Rafindadi, A.A.; Kondo, K.A. 2018. Public finance and rural development in Nigeria: Empirical evidence from the structural equation modelling, *Asian Economic and Financial Review*, 8(11): 1313-1339. <https://doi.org/10.18488/journal.aefr.2018.811.1313.1339>

Refsgaard, K.; Johnson, T.G. 2010. Modelling policies for multifunctional agriculture and rural development-a norwegian case study, *Environmental Policy and Governance*, 20(4): 239-257. <https://doi.org/10.1002/eet.549>

Rushton, G. 1992. Lessons from the debate on location analysis in rural economic development, *International Regional Science Review*, 15(3): 317-324. <https://doi.org/10.1177/016001769301500311>

Salimyanova, I.G.; Novikov, A.A.; Novikova, E.V.; Lushchik, I.V.; Savderova, A.F.; Berezina, N.V.; Rudenko, L.G.; Allalyev, R.M. 2019. Economy digitalization: Information impact on market entities, *Journal of Environmental Treatment Techniques*, 7(4): 654-658.

Saqalli, M.; Gérard, B.; Bielders, C.L.; Defourny, P. 2011. Targeting rural development interventions: Empirical agent-based modeling in Nigerien villages, *Agricultural Systems*, 104(4): 354-364. <https://doi.org/10.1016/j.agsy.2010.12.007>

Yessilov, A. 2015. Social work with immigrants in Kazakhstan as means of prevention of the social conflicts in society, *Mediterranean Journal of Social Sciences*, 6(5): 259-264.

Oleg V. ZAKHARCHENKO

ORCID ID: orcid.org/0000-0001-8198-6569

