Regressive model, Bayesian network, short-term forecasting, GDP of Ukraine

Iryna KARAYUZ[*]*, Petro BIDYUK*[**]

FORECASTING GDP GROWTH RATE IN UKRAINE WITH ALTERNATIVE MODELS

Abstract

The problem of constructing mathematical model for short-term forecasting of GDP is considered. First, extended autoregression is constructed that takes two additional independent variables into consideration. The model resulted provides a possibility for generating short-term forecasts of GDP though not of high quality. Another model was constructed in the form of a Bayesian network. The model turned out to be better than the multiple regression, it provides quite good estimates for probabilities of GDP growth direction.

1. INTODUCTION

Gross domestic product (GDP) is an integrated macroeconomic process that together with level of inflation characterizes to quite acceptable extent current state of macroeconomics. There exists a substantial interest to modeling and forecasting GDP with hiring various model types [1 – 3]. In [1] the issues are analyzed of pooling models for a given set of individual N units observed over T periods of time. It is shown that the forecasting results received by the authors exhibit high quality and outperform ordinary least squares based forecasts. The generalized factor model with infinite dynamics and non-orthogonal idiosyncratic components is considered in [2]. The authors constructed the coincident index for European Union. The dynamic factor model for forecasting the euro area GDP from monthly indicators is developed in [3]. It was found here that surveys and financial data contain important information for the GDP forecasts beyond the monthly real activity measures.

[*] postgraduate student, National Technical University of Ukraine "Kyiv Polytechnic Institute", Institute for Applied System Analysis, 03056, building № 35, 37 Prosp. Peremohy, NTUU "KPI", Kiev, Ukraine, +38 (050) 145-72-54, afanasievaiv@ya.ru

[**] professor, Dr. of Eng. Sci., National Technical University of Ukraine "Kyiv Polytechnic Institute", Institute for Applied System Analysis, 03056, building № 35, 37 Prosp. Peremohy, NTUU "KPI", Kiev, Ukraine, +38 (097) 213-47-08, pbidyuke@gmail.com

As an indicator of inflation process very often is used well known consumer price index (CPI) [4]. To determine correct current estimate of macroeconomic state and to evaluate short-term forecast estimates for economic process of interest it is necessary to construct mathematical models using historical statistical data. At the same time the use of well substantiated mathematical models and forecasts that are based on them does not exclude the possibility of quality expert estimates usage. These estimates could be useful in modeling so that to determine initial conditions or prior probabilities, to form restrictions on key variables of interest, to select appropriate modeling techniques, to estimate elements of model structure, hidden variables etc [5].

The forecasting models for financial and economic processes can be of different complexity level. A wide class of such models may include the following types: multiple regression, autoregression with moving average (ARMA), autoregression with integrated moving average (ARIMA), nonlinear and polynomial regressions as well as numerous modifications and combinations of the structures mentioned. Also highly popular today are the models constructed with the use of artificial intelligence techniques such as neural networks, fuzzy logic, neuro-fuzzy models, a wide class of Bayesian type models (Bayesian regression and Bayesian networks), and the models built with application of support vector machine methodology [6, 7].

2. PROBLEM STATEMENT

The main task of the study is to construct mathematical model describing growth rate for the Ukrainian GDP. Dynamic Bayesian network should be constructed reflecting existing cause-and-effect links between selected key variables and generating short term probabilistic forecasts. Finally a comparison of forecasts estimates should be performed computed with different model types.

3. ANALYSIS OF CURRENT SITUATION IN UKRAINE

Using the economic production methodology GDP could be found as a sum of gross value added (GVA) for all kinds of economic activities plus net tax [8]. The model should take into consideration growth rates for economy branches, consumer price index (CPI), production price index (PPI), GVA structure changes etc.

Actual income of population has decreased in 2014 by 8.4%; population's savings decreased by 26.5%; CPI increased by 24.9%; PPI for industrial production increased by 17.1% [9].

The dynamics of structure of goods and services production for separate economic activities is changing as shown in Fig. 1 [9]. The GVA structure is influenced by the following factors:

- changes of law (for example, rate of rent for natural gas mining for joint ventures is 70% from the cost of final product [9] what negatively influences attractiveness of such activities in 2015);
- local war (LW) in the south-east of Ukraine (Lugansk and Donetsk areas) has highly negatively influenced the volume of industrial production of these areas and Ukraine as a whole due to existing links between enterprises;
- temporary occupation of Autonomous Crimean Republic, Sebastopol area.

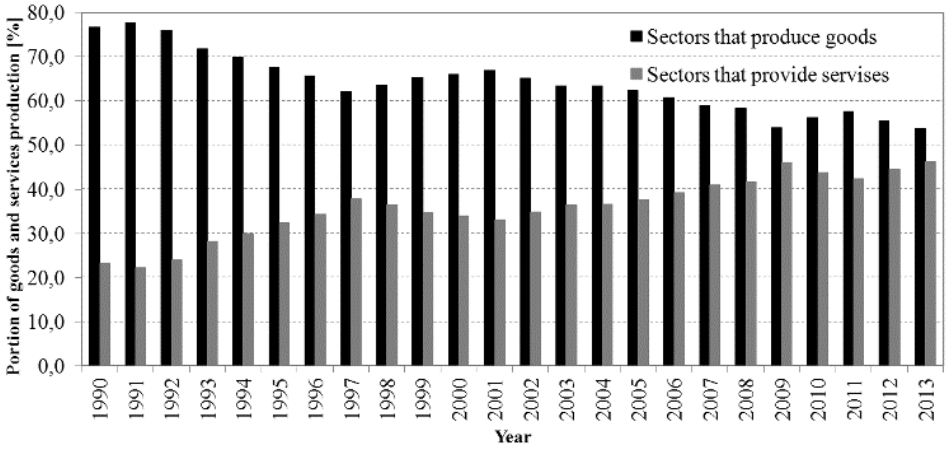


Fig. 1. Dynamics of the structure of goods and services production [source: own study]

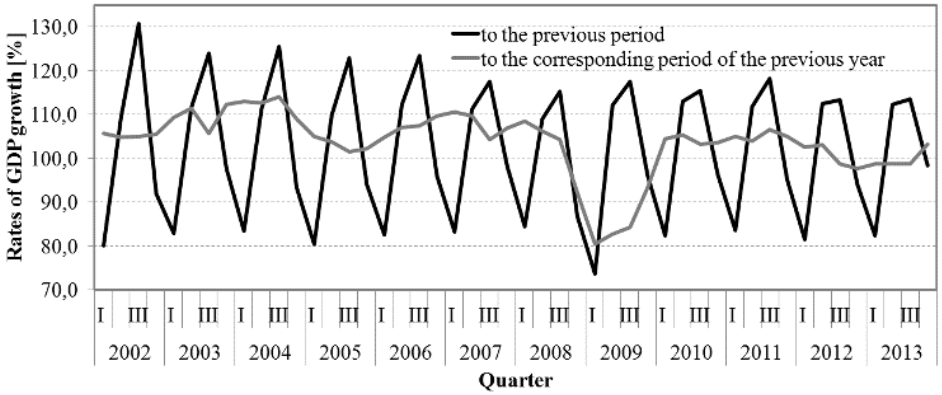


Fig. 2. Rates of GDP growth, quarterly data [source: own study]

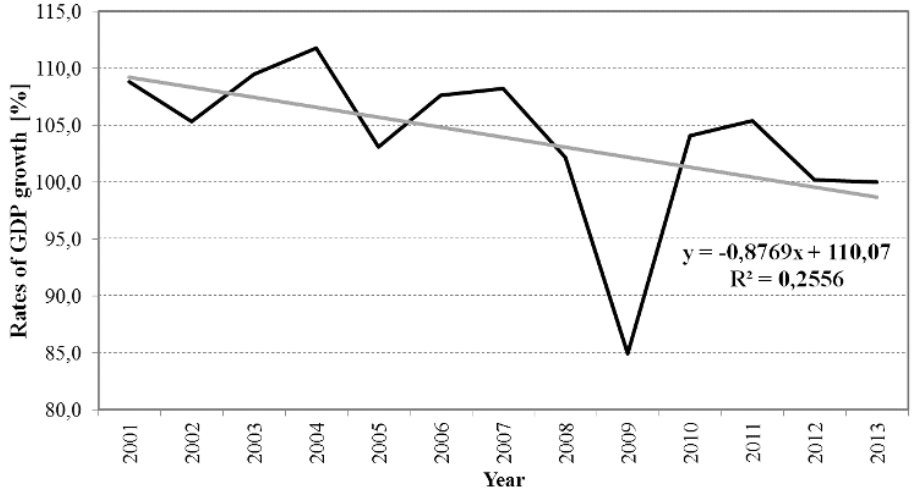


Fig. 3. Rates of GDP growth GDP, annual data [source: own study]

Figs. 2 and 3 rates of growth for GDP with respect to the same period of previous year. Each factor is informative for modeling and to reach quality forecasting it is necessary to take into consideration all the time series mentioned.

4. REGRESSION MODEL

First, consider the possibility for constructing model in the form of extended autoregression (ARX):

$$y(k) = a_0 + \sum_{i=1}^{p} a_i y(k-i) + \sum_{j=1}^{q} b_j x(k-j) + \varepsilon(k),$$

where: p – autoregression order,

q – a number of regressors,

$y(k)$ – the main dependent variable at discrete moment of time k,

$x(k)$ – regressor,

a_i – i-th autoregression part parameter,

b_j – j-th parameter of multiple regression part of the model,

$\varepsilon(k)$ – random process that is formed by nonmeasurable stochastic external disturbances, measurement errors, parameter computing errors, and model structure inadequacy.

Such models provide a possibility for constructing forecasting functions that enable computing of forecasts for a necessary number of steps using conditional mathematical expectation operator.

According to the correlation analysis performed the model should include CPI and PPI with lags 3 and 2, respectively:

$$GDP(k) = 75{,}8 + 0{,}94 \cdot GDP(k-4) - 0{,}21 \cdot GDP\_4(k-4) - \\ - 0{,}15 \cdot CPI(k-3) - 0{,}31 \cdot PPI(k-2).$$

where: GDP is relative GDP growth to the previous time period, %,
GDP_4 is relative GDP growth to the same time period of previous year, %,
CPI is CPI growth with respect to the previous time period,%,
PPI is PPI growth with respect to the previous time period,%.

Adequacy of this model is characterised by the following statistics: $R^2 = 0{,}95$; $DW = 1{,}55$; $SSE = 429{,}1$. The parameters estimates correspond to the realities of GDP changes in time: when CPI and PPI are growing GDP growth rate is decreasing.

The forecasts quality is characterized by the following statistics: MSE = 389,7; MAPE = 16,8%; U = 0,13 (Theil coefficient).

Tab. 1. Comparison of actual GDP (SSSU) and its forecast for 2014

Time period	GDP, %		Forecast error, %
	SSSU	Model	
QI 2014	**76,4**	83,2	8,91
QII 2014	**108,4**	111,1	2,52
QIII 2014	**112,0**	110,8	-1,03
QIV 2014	**88,8**	91,8	3,34
2014 year	**93,2**	**100,2**	**7,55**

The statistical data in Table 1 were taken from the State Statistical Service of Ukraine (SSSU) as of March 20, 2015 [9].

As far as mean absolute percentage error is about 17% for the regression ARX model considered above it is necessary to try another approach to modeling GDP that would be ideologically different. Besides, this model cannot describe quantitatively and qualitatively influence of the war in the south-east of Ukraine. Also one of important influence factors to development of GDP is support of the dollar/hryvna exchange rate practically at fixed level) from May 2012 till February 2014. Define this event as “Pseudo Stability).

5. BAYESIAN NETWORK CONSTRUCTION

Bayesian network (BN) is a directed acyclic graph vertices of which are model variables, and its arcs show existing cause-and-effect links between the variables. To construct a model in the form of a Bayesian network introduce the following variables: Local War (LW) is a variable that takes value of "1" when there is a war on Ukrainian territory (southeast); Pseudo Stability (PS) takes a value of "1" if exchange rate for hryvna (Ukrainian currency) is fixed; Investment Climate (IC) takes value of "1" if the rate of investments growth into Ukraine is positive; CPI(-3) and PPI(-3) are equal "1" if their growth rate exceeds 15%; GDP is equal "1" if its growth rate is positive.

The first BN was built by hiring expert estimates and with the use of model quality criterion called minimum description length (MDL), it is shown n Fig. 4. The models given in Fig. 5 were constructed by heuristic algorithm [10] with the use of computed values for mutual information in the following way: a) without restrictions; b) with restriction that variable 6.GDP can be only subordinated (child), and the variables "1. LW", "2. PS" can only be of parents type; c) with restrictions of type b) and without free vertices.

The net given in Fig. 5a cannot be used practically because GDP growth rate cannot influence growth rate of PPI that was three quarters (9 months) before.

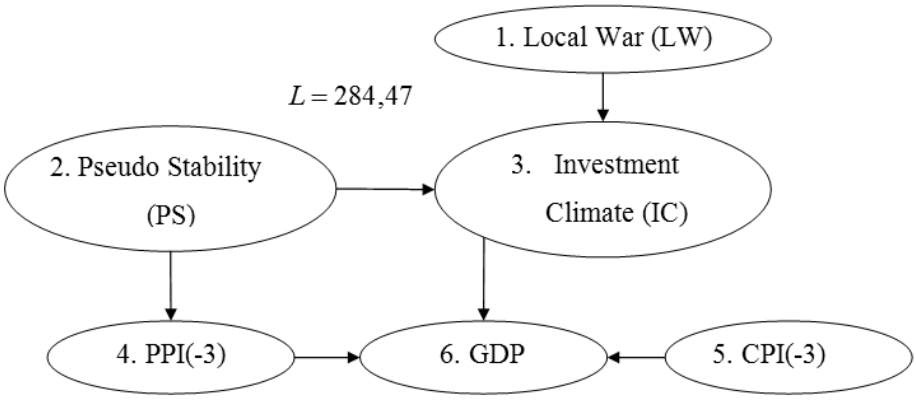


Fig. 4. Network for GDP of Ukraine built by expert [source: own study]

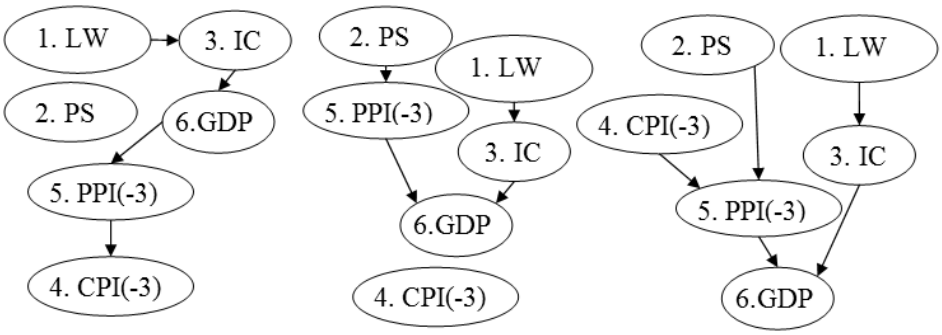


Fig. 5. Networks for GDP of Ukraine built by heuristic algorithm: a) without restrictions; b) with restriction that variable 6.GDP can be only subordinated (child), and the variables "1. LW", "2. PS" can only be of parents type; c) with restrictions of type b) and without free vertices [source: own study]

P(PS)	
PS=0	PS=1
0,85	0,15

P(LW)	
LW=0	LW=1
0,25	0,75

	P(PPI(-3)/PS)	
	PPI(-3)=0	PPI(-3)=1
PS=0	0,49	0,51
PS=1	1	0

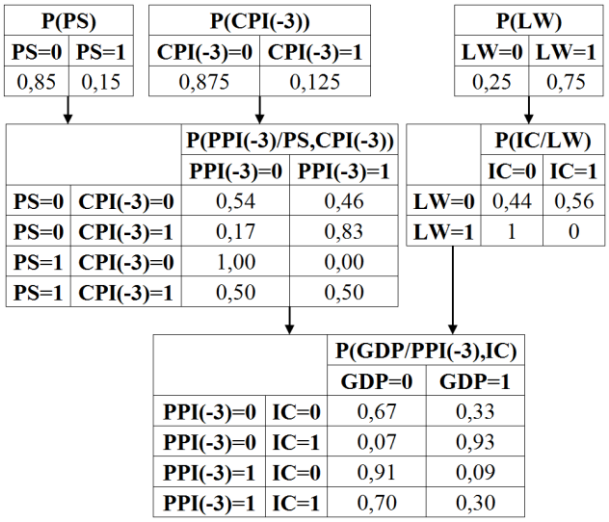
	P(IC/LW)	
	IC=0	IC=1
LW=0	0,44	0,56
LW=1	1	0

		P(GDP/PPI(-3),IC)	
		GDP=0	GDP=1
PPI(-3)=0	IC=0	0,67	0,33
PPI(-3)=0	IC=1	0,07	0,93
PPI(-3)=1	IC=0	0,91	0,09
PPI(-3)=1	IC=1	0,70	0,30

Fig. 6. Conditional probability tables for BN in Fig. 5b [source: own study]

Expression for computing joint probability function for BN in Fig. 6:

$$P(PS, LW, PPI(-3), IC, GDP) = \sum \left\{ P(GDP \mid PPI(-3), IC) \cdot \right.$$
$$\left. \cdot \sum \left\{ P(PPI(-3) \mid PS) \cdot P(PS) \right\} \sum \left\{ P(IC \mid LW) \cdot P(LW) \right\} \right\}$$



P(PS)	
PS=0	PS=1
0,85	0,15

P(CPI(-3))	
CPI(-3)=0	CPI(-3)=1
0,875	0,125

P(LW)	
LW=0	LW=1
0,25	0,75

		P(PPI(-3)/PS,CPI(-3))	
		PPI(-3)=0	PPI(-3)=1
PS=0	CPI(-3)=0	0,54	0,46
PS=0	CPI(-3)=1	0,17	0,83
PS=1	CPI(-3)=0	1,00	0,00
PS=1	CPI(-3)=1	0,50	0,50

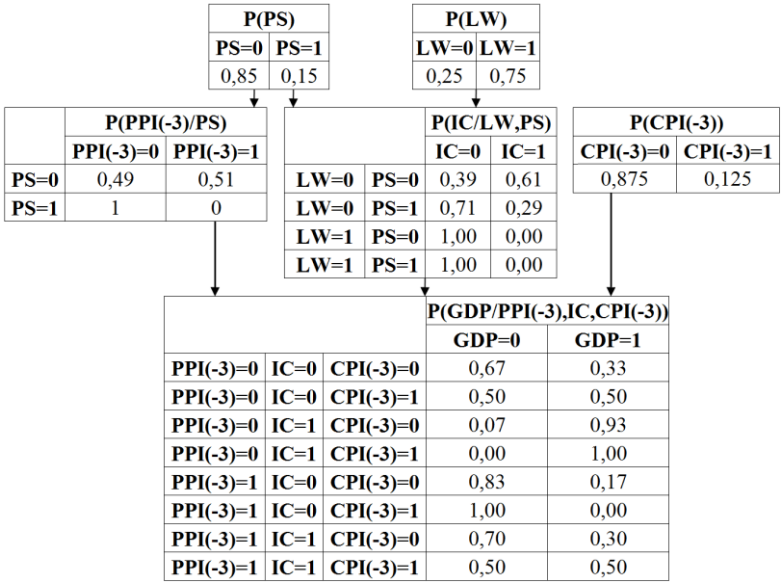
	P(IC/LW)	
	IC=0	IC=1
LW=0	0,44	0,56
LW=1	1	0

		P(GDP/PPI(-3),IC)	
		GDP=0	GDP=1
PPI(-3)=0	IC=0	0,67	0,33
PPI(-3)=0	IC=1	0,07	0,93
PPI(-3)=1	IC=0	0,91	0,09
PPI(-3)=1	IC=1	0,70	0,30

Fig. 7. Conditional probability tables for BN in Fig. 5c [source: own study]

Expression for computing joint probability function for BN in Fig. 7:

$$P(PS,CPI(-3),LW,PPI(-3),IC,GDP)=\sum\left\{P(GDP|PPI(-3),IC)\cdot\right.$$
$$\left.\cdot\sum\left\{P(PPI(-3)|PS,CPI(-3))\cdot P(PS)\cdot P(CPI(-3))\right\}\sum\left\{P(IC|LW)\cdot P(LW)\right\}\right\}$$



P(PS)	
PS=0	PS=1
0,85	0,15

P(LW)	
LW=0	LW=1
0,25	0,75

	P(PPI(-3)/PS)	
	PPI(-3)=0	PPI(-3)=1
PS=0	0,49	0,51
PS=1	1	0

		P(IC/LW,PS)	
		IC=0	IC=1
LW=0	PS=0	0,39	0,61
LW=0	PS=1	0,71	0,29
LW=1	PS=0	1,00	0,00
LW=1	PS=1	1,00	0,00

P(CPI(-3))	
CPI(-3)=0	CPI(-3)=1
0,875	0,125

			P(GDP/PPI(-3),IC,CPI(-3))	
			GDP=0	GDP=1
PPI(-3)=0	IC=0	CPI(-3)=0	0,67	0,33
PPI(-3)=0	IC=0	CPI(-3)=1	0,50	0,50
PPI(-3)=0	IC=1	CPI(-3)=0	0,07	0,93
PPI(-3)=0	IC=1	CPI(-3)=1	0,00	1,00
PPI(-3)=1	IC=0	CPI(-3)=0	0,83	0,17
PPI(-3)=1	IC=0	CPI(-3)=1	1,00	0,00
PPI(-3)=1	IC=1	CPI(-3)=0	0,70	0,30
PPI(-3)=1	IC=1	CPI(-3)=1	0,50	0,50

Fig. 8. BN with nodes as conditional probability tables for BN on fig. 4 [source: own study]

Expression for computing joint probability function for BN in Fig. 8:

$$P(PS, CPI(-3), LW, PPI(-3), IC, GDP) =$$
$$= \sum \left\{ P(GDP \mid PPI(-3), IC, CPI(-3)) \cdot \sum \left\{ P(IC \mid LW, PS) \cdot P(LW) \cdot P(PS) \right\} \cdot \right.$$
$$\left. \cdot \sum \left\{ P(PPI(-3) \mid PS) \cdot P(PS) \right\} \cdot P(CPI(-3)) \right\}$$

Using the joint probability function and Bayes' theorem we got the following CPTs:

Tab. 2. Conditional probabilities for constructing BNs

	BN of Fig. 6b		BN of Fig. 6c		BN of Fig. 5	
	P(GDP/LW)		**P(GDP/LW)**		**P(GDP/LW)**	
	GDP=0	**GDP=1**	**GDP=0**	**GDP=1**	**GDP=0**	**GDP=1**
LW=0	0,534	0,466	0,535	0,465	0,510	0,490
LW=1	0,833	0,167	0,833	0,167	0,737	0,263

Thus, the least possible probabilistic estimate for the negative direction of Ukrainian GDP growth in 2014 is 0.74 by condition that takes place local war. Respectively the most possible probabilistic estimate for the positive direction of Ukrainian GDP growth in 2014 is 0.26.

6. CONCLUSIONS

It was shown that application of regression models for describing the processes of economy in transition not always finalizes with positive results. This can be explained by numerous out-of-market events (factors) that influence development of the economy. The short-term forecasting results obtained in this case are not of very high quality though acceptable.

On the other side probabilistic models such as Bayesian networks provide a possibility for obtaining well substantiated and quite good probabilistic estimates for the direction of GDP growth in Ukraine. A substantial advantage of the simple heuristic method used for constructing BN is in its transparency and a small number of operations. Disadvantage of the heuristic approach is that after every increase in a number of variables it is necessary to reconstruct the model graph (model structure). The minimum description length based algorithm operates according to the principle of entropy minimization (the higher is dependence between the variables the better is the case from the modeling point of view). Although such approach may result in some paradoxes regarding final model quality.

In the future we will come to the new modeling problem for quantitative estimation of Ukrainian GDP growth rate: the regression model constructed in the paper should be modified with the results of structural analysis for GVA, and the interest rate of National Bank should also be taken into consideration.

REFERENCES

[1] HOOGSTRATE A., PALM F., PFANN G.: *Pooling in Dynamic Panel-Data Models: An Application to Forecasting GDP Growth Rates.* Journal of Business & Economic Statistics, 2000, Vol. 18, No. 3, pp. 274-283.

[2] FORNI M., HALLIN M., LIPPI M., REICHLIN L.: *The Generalized Dynamic-Factor Model: Identification and Estimation*. The Review of Economics and Statistics, 2000, Vol. 82, No. 4, pp. 540-554.

[3] BANBURA M., RÜNSTLER G.: *A Look into the Factor Model Black Box: Publication Lags and the Role of Hard and Soft Data in Forecasting GDP*. International Journal of Forecasting, 2011, Vol. 27, No. 2, pp. 333-346.

[4] BRANSON W.: *Macroeconomic theory and policy.* New York: Harper & Row Publishers, 1989, p. 656.

[5] DELURGIO S.A.: *Forecasting principles and applications*. McGraw-Hill, Boston 1998, p. 802.

[6] RUTKOWSKI L.: *Metody i techniki sztucznej inteligencji.* Wydawnictwo Naukowe PWN, Warszawa 2005, p. 520.

[7] POLE A., WEST M., HARRISON J.: *Applied Bayesian forecasting and time series analysis.* Chapman & Hall/CRC, New York 1994, p. 409.

[8] Macroeconomics / Ed. by V.D. Basilevych, Znannya, Kyiv 2007, p. 703.

[9] The State Statistical Service of Ukraine: http://www.ukrstat.gov.ua

[10] ZGUROWSKY M.Z., BIDYUK P.I., TERENTYEV O.M., PROSYANKINA T.I.: *Bayesian networks and decision trees.* Edelveis, Kyiv 2015, p. 360.