Forecasting of Cereal Crop Harvest on the Basis of an Extrapolation Canonical Model of a Vector Random Sequence

Igor Atamanyuk[1], Yuriy Kondratenko[2], Anastasiya Poltorak[1], Natalia Sirenko[1], Vyacheslav Shebanin[1], Inna Baryshevska[1], Valeriia Atamaniuk[3]

[1] Mykolayiv National Agrarian University, Georgiya Gongadze Str. 9, 54010 Mykolaiv, Ukraine
atamanyuk@mnau.edu.ua, sirenko@mnau.edu.ua, shebanin@mnau.edu.ua, BaryshevskaIV@mnau.edu.ua, poltorak@mnau.edu.ua
[2] Petro Mohyla Black Sea National University, 68th Desantnykiv Str. 10, 54003 Mykolaiv, Ukraine
y_kondrat2002@yahoo.com
[3] Western Norway University of Applied Sciences, Bjørnsons Str. 45, 5528 Haugesund, Norway
atamanyuk.le@gmail.com

Abstract. The work is devoted to the solving of an important economic problem of the forecasting of cereal crop harvest. A stochastic character of the change of crop yield figures because of the influence of random weather-related factors is an essential peculiarity of this problem. Therefore, to forecast the cereal crop harvest, the methods of random sequence analysis are proposed to use. The developed extrapolation method doesn't impose any restrictions on a forecasted random sequence of the change of crop yield figures (linearity, stationarity, Markov behavior, monotony, etc.). Taking into full account stochastic peculiarities of the conditions of cereal crop production and crop yield figures allows to achieve maximum accuracy of a forecasting problem solving. The block diagram of an algorithm introduced in the work represents the peculiarities of the calculation of the predictive model parameters. The expression for calculation of an extrapolation error allows to determine necessary volume of a priori and a posteriori information for achieving required quality of a forecasting problem solving. The results of a numerical experiment confirmed high efficiency of the suggested method of forecasting of the cereal crop harvest.

Keywords: calculation method, random sequence, canonical decomposition, prognostication of the crop.

CEUR-WS.org/Vol-2393/paper_276.pdf

1 Introduction

Volume of grain production essentially influences standards of living, contributes to food security of a country. Crop capacity deserves special attention among many figures characterizing the level of cereal crops production. Solving of the problems of formation of food reserve funds, availability of necessary facilities for storing of obtained harvest, forming of adequate and efficient foreign trade policy greatly depend on the accuracy of its forecasting. The actuality of crop yield forecasting gets special importance when developing management decisions under conditions of uncertainty including the conditions of economic instability. In the light of the issues of food security, the problem of crop yield forecasting is actual not only for Ukraine as one of the biggest grain producers but also for international community [1,2].

Crop productivity is a complex figure from the point of view of the forecasting because harvest formation is connected with the influence of both production factors and weather conditions and also greatly depends on the peculiarities of biological systems [2].

At present different approaches to the crop yield forecasting are developed and applied in practice on the basis of:

1) analysis of trend and cyclicality in the crop yield dynamics [3,4];
2) identification of the analogous year [5];
3) forming of regression dependencies between different statistical data and data obtained on the basis of remote and meteorological observations [4];
4) modeling [4,6];
5) analysis of synoptic processes [7].

Approaches of the first, second and fifth groups distinguish with great lead time as well as with insufficient accuracy. The approaches of the third and the fourth groups are most widely used. In most cases meteorological data are used as input information for building regression or for the modeling of the processes of plant growth. In these cases the forecasting is based mostly on the use of indirect factors, not on the analysis of the actual state of soil and peculiarities of the use of fertilizers for plant nutrition and soil fertility improvement.

Dynamic models [4] used today don't take into account the whole background of the change of crop yield figures and conditions of cereal crop production which significantly restricts the accuracy of such models.

A main peculiarity of crop capacity is a stochastic character of the change of this figure. Thereupon, to forecast the cereal crop harvest for the purpose of maximum use of the production background and to take into a full account the influence of different random factors (amount of precipitation, air and soil temperature, number of sunny days, humidity, etc.), it is necessary to use the methods and algorithms of the theory of random functions and random sequences.

The aim of this work is the development of the efficient and robust method for forecasting of the cereal crop harvest. The main requirement to the forecasting method is the absence of any essential limitations on the stochastic properties of the accidental process of change of the cereal crop harvest.

2 Related Works and Problem Statement

Methods of artificial intelligence that are used for the forecasting of random sequences have restricted accuracy characteristics and are applied as a rule in case of small volume of statistic data [8,9]. When analysing the cereal crop harvest, quite large volume of information can be accumulated at expense of increasing of data detailing (figure concretization at the regional level and agricultural enterprises; monthly accounting for temperature, moisture, quantity of fertilizers; use of soil characteristics, etc.). Therefore, to formulate mathematical models, it is expedient to apply deductive methods of forecasting on the basis of maximum volume of a priori information. Kolmogorov-Garbor polynomial [10] is the most general extrapolation form to solve the problem of non-linear extrapolation, but determination of its parameters for a large number of known values and used order of non-linear relations is a very difficult and laborious procedure (thus, for 11 known values and 4th order nonlinearity, it is necessary to obtain and solve 1819 equations of partial derivatives of mean-square error of extrapolation). Thereupon, when forming realizable in practice algorithms of the forecasting, different simplifications and restrictions on the properties of a random sequence are used. For example, a number of suboptimal methods [11] of non-linear extrapolation with a bounded order of a stochastic relation on the basis of approximation of a posteriori density of probabilities of an estimated vector by orthogonal Hermite polynomial expansion or in the form of Edgeworth series is offered by V.S. Pugachev. The solution of non-stationary A. N. Kolmogorov equation (a particular case of R. L. Stratanovich differential equation [12] for description of Markovian processes) is obtained provided that a drift coefficient is a linear function of the state, and a diffusion coefficient equals to a constant. An exhaustive solution of the problem of optimal linear extrapolation for different classes of random sequences and different level of informational support of a forecasting problem exists (A.N. Kolmogorov equation for stationary random sequences measured without errors; Kalman method [13] for Markov noisy random sequences; Wiener-Hopf filter-extrapolator [14] for noisy stationary sequences; algorithms of optimal linear extrapolation of V.D. Kudritsky [15] on the basis of linear canonical expansion of V.S. Pugachev, etc.). However, maximum accuracy of the forecasting with the help of the methods of linear extrapolation can be achieved only for Gaussian random sequences. Forecasting method [16,17] on the basis of non-linear canonical expansion is the most universal with regard to limitations (linearity, Markov property, stationarity, monotony, scalarity, etc.) imposed on the properties of the sequences of random values. Application of this method will allow to take full account of peculiarities of the change of cereal crop harvest and, consequently, to achieve maximum quality of forecasting.

3 Theoretical Conception of the Proposed Forecasting Method

Vector random sequence $\overline{X}(i)=\{X_h(i)\},\ i=\overline{1,I},\ h=\overline{1,H}$ is to be considered. Components are random sequences describing the change of the crop yield figures of certain cereal crops (wheat, rye, barley, etc.), the change of natural conditions (temperature, precipitation amount, number of sunny days, etc.) and also intensity of the use of mineral and organic fertilizers at discrete points of time $t(i)$ (as a rule with discrete step which is equal to one year for mesoeconomic and macroeconomic forecasting).

Non-linear canonical expansion of a vector random sequence can be written as [18,19]:

$$X_h(i) = M\left[X_h(i)\right] + \sum_{\nu=1}^{i-1}\sum_{l=1}^{H}\sum_{\lambda=1}^{N} W_{\nu l}^{(\lambda)} \beta_{l\lambda}^{(h,1)}(\nu,i) + \sum_{l=1}^{h-1}\sum_{\lambda=1}^{N} W_{il}^{(\lambda)} \beta_{l\lambda}^{(h,1)}(i,i) + \\ + W_{ih}^{(1)},\ i=\overline{1,I}. \tag{1}$$

Discretized moment functions $M\left[X_l^{\lambda}(\nu)\right]$, $M\left[X_l^{\lambda}(\nu)X_h^{s}(i)\right]$, $\nu,i=\overline{1,I}$; $l,h=\overline{1,H}$; $\lambda,s=\overline{1,N}$ are source information for the model of a random sequence.

Random coefficients $D_{l,\lambda}(\nu),\ l=\overline{1,H},\ \lambda=\overline{1,N},\ \nu=\overline{1,I}$ and non-random coordinate functions, $\beta_{l\lambda}^{(h,s)}(\nu,i)$, l,h $=\overline{1,H},\ \lambda,s=\overline{1,N},\ \nu,i=\overline{1,I}$ are determined with the help of following expressions (algorithm of parameter calculation is presented in Fig.1):

$$W_{\nu l}^{(\lambda)} = X_l^{\lambda}(\nu) - M\left[X_l^{\lambda}(\nu)\right] - \sum_{\mu=1}^{\nu-1}\sum_{m=1}^{H}\sum_{j=1}^{N} W_{\mu m}^{(j)} \beta_{mj}^{(l,\lambda)}(\mu,\nu) - \\ - \sum_{m=1}^{l-1}\sum_{j=1}^{N} W_{\nu m}^{(j)} \beta_{mj}^{(l,\lambda)}(\nu,\nu) - \sum_{j=1}^{\lambda-1} W_{\nu l}^{(j)} \beta_{lj}^{(l,\lambda)}(\nu,\nu),\ \nu=\overline{1,I}; \tag{2}$$

$$D_{l,\lambda}(\nu) = M\left[\left\{W_{\nu l}^{(\lambda)}\right\}^2\right] = M\left[X_l^{2\lambda}(\nu)\right] - M^2\left[X_l^{\lambda}(\nu)\right] - \\ - \sum_{\mu=1}^{\nu-1}\sum_{m=1}^{H}\sum_{j=1}^{N} D_{mj}(\mu)\left\{\beta_{mj}^{(l,\lambda)}(\mu,\nu)\right\}^2 - \sum_{m=1}^{l-1}\sum_{j=1}^{N} D_{mj}(\nu)\left\{\beta_{mj}^{(l,\lambda)}(\nu,\nu)\right\}^2 - \\ - \sum_{j=1}^{\lambda-1} D_{lj}(\nu)\left\{\beta_{lj}^{(l,\lambda)}(\nu,\nu)\right\}^2,\ \nu=\overline{1,I}; \tag{3}$$

$$\beta_{l\lambda}^{(h,s)}(\nu,i) = \frac{M\left[W_{\nu l}^{(\lambda)}\left(X_h^{s}(i) - M[X_h^{s}(i)]\right)\right]}{M\left[\left\{W_{\nu l}^{(\lambda)}\right\}^2\right]} = \\ = \frac{1}{D_{l\lambda}(\nu)}(M\left[X_l^{\lambda}(\nu)X_h^{s}(i)\right] - M\left[X_l^{\lambda}(\nu)\right]M\left[X_h^{s}(i)\right] - \tag{4}$$

$$-\sum_{\mu=1}^{\nu-1}\sum_{m=1}^{H}\sum_{j=1}^{N}D_{mj}\left(\mu\right)\beta_{mj}^{(l,\lambda)}\left(\mu,\nu\right)\beta_{mj}^{(h,s)}\left(\mu,i\right)-$$

$$-\sum_{m=1}^{l-1}\sum_{j=1}^{N}D_{mj}(\nu)\beta_{mj}^{(l,\lambda)}\left(\nu,\nu\right)\beta_{mj}^{(h,s)}\left(\nu,i\right)-$$

$$-\sum_{j=1}^{\lambda-1}D_{lj}(\nu)\beta_{lj}^{(l,\lambda)}\left(\nu,\nu\right)\beta_{lj}^{(h,s)}\left(\nu,i\right),\ \lambda=\overline{1,h},\ \nu=\overline{1,i}.$$

Random sequence $X_h\left(i\right),\ h=\overline{1,H},\ i=\overline{1,I}$ is represented with the help of $H\times N$ arrays $\left\{W_l^{(\lambda)}\right\},\ l=\overline{1,H},\ \lambda=\overline{1,N}$ of uncorrelated centered random coefficients $W_{\nu l}^{(\lambda)},\ \nu=\overline{1,I}$. Each of these coefficients contains information about the corresponding value $X_l^{\lambda}\left(\nu\right)$ (crop yield figures of cereal crops, precipitation amount, intensity of the use of mineral and organic fertilizers, etc.) and coordinate functions $\beta_{l\lambda}^{(h,s)}\left(\nu,i\right)$ describe probabilistic relations of $\lambda+s$ order between components $X_l\left(\nu\right)$ and $X_h\left(i\right)$ (the impact of various factors on the crop).

Expression (1) is also true if some stochastic relations of a random sequence $\overline{X}\left(i\right)=\left\{X_h\left(i\right)\right\}$ are absent. In this case the corresponding coordinate functions take value 0 and these relations are automatically excluded from a canonical expansion.

Vector algorithm of extrapolation for arbitrary number of components $X_h\left(i\right),\ h=\overline{1,H};\ i=\overline{1,I}$ and N order of stochastic relations on the basis of a canonical expansion (1) is of the form [20]:

$$m_{j,h}^{(\mu,l)}\left(s,i\right)=\begin{cases}M\left[X\left(i\right)\right],\ \text{if } \mu=0,\\ m_{j,h}^{(\mu,l-1)}\left(s,i\right)+\left(x_j^l\left(\mu\right)-m_{j,j}^{(\mu,l-1)}\left(l,\mu\right)\right)\beta_{j,l}^{(h,s)}\left(\mu,i\right),\\ \text{if } l>1,\ j>1,\\ m_{j-1,h}^{(\mu,N)}\left(s,i\right)+\left(x_j^1\left(\mu\right)-m_{j-1,j}^{(\mu,N)}\left(1,\mu\right)\right)\beta_{j-1,N}^{(h,s)}\left(\mu,i\right),\\ \text{if } l=1,\ j>1,\\ m_{H,h}^{(\mu-1,N)}\left(s,i\right)+\left(x_1^1\left(\mu\right)-m_{H,1}^{(\mu-1,N)}\left(1,\mu\right)\right)\beta_{1,1}^{(h,s)}\left(\mu+1,i\right),\\ \text{for } l=1,\ j=1.\end{cases}\tag{5}$$

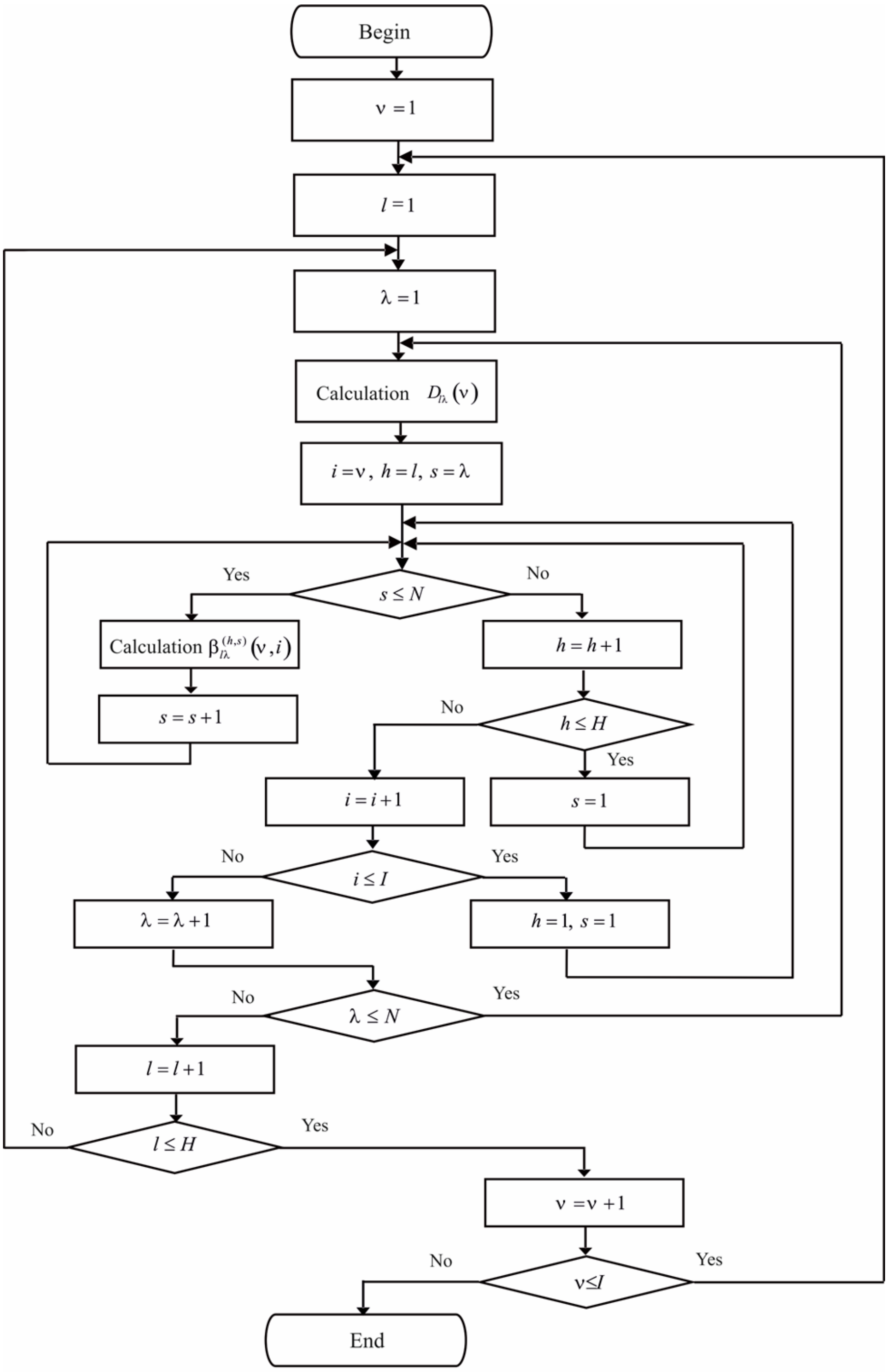


Fig. 1. Block-diagram of the algorithm for calculation of model (1) parameters

where

$m_{x;j,h}^{(\mu,l)}(1,i) = M\left[X_h(i)/x_\lambda^n(\nu),\ \lambda=\overline{1,H}, n=\overline{1,N}, \nu=\overline{1,\mu-1};\ x_\lambda^n(\mu),\ \lambda=\overline{1,j}, n=\overline{1,l}\right]$ - is optimal in mean-square sense estimation of future values of an investigated random sequence provided that a posteriori information $x_\lambda^n(\nu), \lambda=\overline{1,H},\ \ n=\overline{1,N}$, $\nu=\overline{1,\mu-1};\ x_\lambda^n(\mu),\ \lambda=\overline{1,j},\ n=\overline{1,l}$ is used for forecasting.

Expression for mean-square error of extrapolation with the help of algorithm (5) by known values $x_j^n(\mu),\ \mu=\overline{1,k}; j=\overline{1,H};\ n=\overline{1,N}$ can be written as

$$E_h^{(k,N)}(i) = M\left[X_h^2(i)\right] - M^2\left[X_h(i)\right] - \sum_{\mu=1}^{k}\sum_{j=1}^{H}\sum_{n=1}^{N} D_{jn}(\mu)\left\{\beta_{jn}^{(h,1)}(\mu,i)\right\}^2,\ i=\overline{k+1,I}. \quad (6)$$

Mean-square error of extrapolation $E_h^{(k,N)}(i)$ equals to the dispersion of a posteriori random sequence

$$X_h^{(k,N)}(i) = X\left(i/x_l^\nu(j),\ \nu=\overline{1,N}, j=\overline{1,k}, l=\overline{1,H}\right) = m_{H,h}^{(k,N)}(1,i) + \sum_{\nu=k+1}^{i-1}\sum_{l=1}^{H}\sum_{\lambda=1}^{N} W_{\nu l}^{(\lambda)}\beta_{l\lambda}^{(h,1)}(\nu,i) + \sum_{l=1}^{h-1}\sum_{\lambda=1}^{N} W_{il}^{(\lambda)}\beta_{l\lambda}^{(h,1)}(i,i) + W_{ih}^{(1)},\ i=\overline{k+1,I}. \quad (7)$$

Calculation method of the forecasting of future values of crop yield figures on the basis of a predictive model (5) involves the realization of the following stages:

Stage 1. Gathering of statistical data on the results of cereal crop harvest and production conditions;

Stage 2. Estimation of moment functions $M\left[X_l^\lambda(\nu)X_h^s(i)\right]$ on the basis of accumulated realizations of a random sequence describing the process of the change of cereal crop harvest;

Stage 3. Calculation of the parameters of extrapolation algorithm (5) with the help of expressions (2)-(4);

Stage 4. Estimation of the quality of solving of the forecasting problem for an investigated sequence using expression (6).

4 Discussion of the Numerical Experiment Results

Forecasting method is approbated on the basis of crop yield data [1,21] of twenty-four regions of Ukraine during the period 2007-2018 (graphs of the change of mathematical expectation and mean-square deviation are presented in Fig. 2). During the process of a numerical experiment a vector random sequence $X_h(i),\ h=\overline{1,5};\ i=\overline{1,12}$ ($X_1(i),\ i=\overline{1,12}$ - wheat productivity, centner/ha; $X_2(i),\ i=\overline{1,12}$ - barley productivity, centner/ha; $X_3(i),\ i=\overline{1,12}$ - humus content, %; $X_4(i),\ i=\overline{1,12}$ - amount of precipitation, mm; $X_5(i),\ i=\overline{1,12}$ - use of mineral fertilizers, kg/ha) was studied.

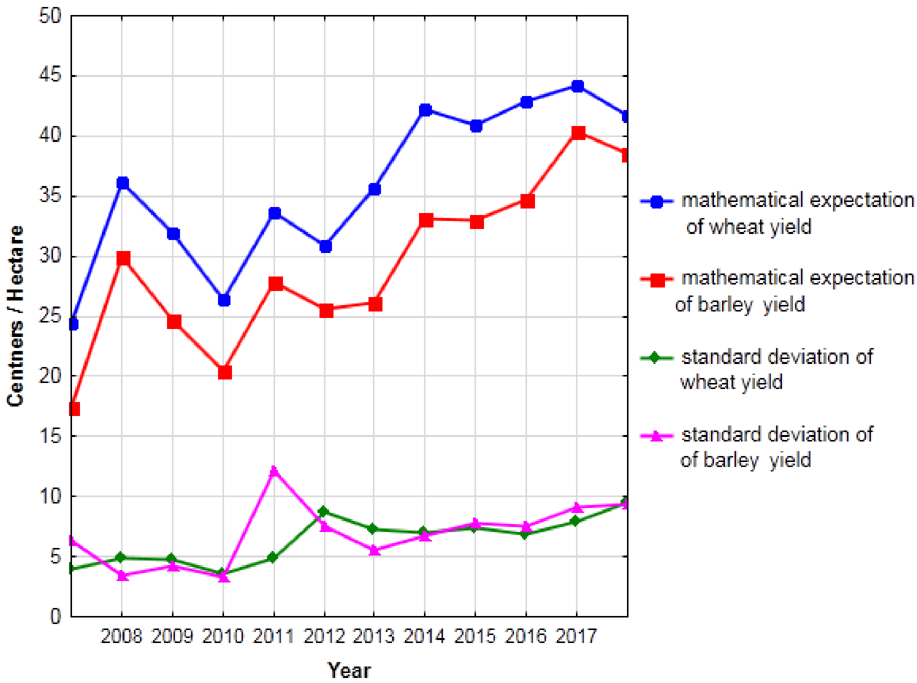


Fig. 2. Characteristics of wheat and barley productivity in Ukraine (2007-2018)

Preliminary investigations on the basis of statistic information showed that stochastic relations of ≤ 4 order are the most sustainable and significant. Thus, 165 values $x_h^{\lambda}(i), h=\overline{1,5},\ i=\overline{1,11},\ \lambda=\overline{1,3}$ and 5220 not equal to zero weight coefficients $\beta_{l\lambda}^{(h,s)}(\nu,i),\ \nu,i=\overline{1,12},\ l,h=\overline{1,5},\ \lambda,s=\overline{1,3}$ were used to forecast the crop yield figures for the last year (2018) in a forecasting algorithm (5). At the initial stage of a numerical experiment moment functions $M\left[X_l^{\lambda}(\nu)X_h^{s}(i)\right],\ \nu,i=\overline{1,12},\ l,h=\overline{1,5},$ $\lambda,s=\overline{1,3}$ were determined, and parameters $\beta_{l\lambda}^{(h,s)}(\nu,i),\ \nu,i=\overline{1,12},\ l,h=\overline{1,5},$ $\lambda,s=\overline{1,3}$ of a predictive model (5) were calculated on that base (experimental investigations were made using software product Fig. 3 that was created in Delphi programming system).

For example, values of autocorrelated functions $M\left[\overset{o}{X}_h(\nu)\overset{o}{X}_h(i)\right],$ $\nu=\overline{1,12},\ i=\overline{1,12},\ h=\overline{1,2}$ for components $X_1(i)$ and $X_2(i),\ i=\overline{1,12}$ are presented in Table 1, Table 2.

For the period 2007-2017 values of autocorrelative functions are calculated by processing of statistic data (crop yield figures in 2007-2017). For 2018 values

Table 1. Autocorrelative function of the component $X_1(i),\ i=\overline{1,12}$

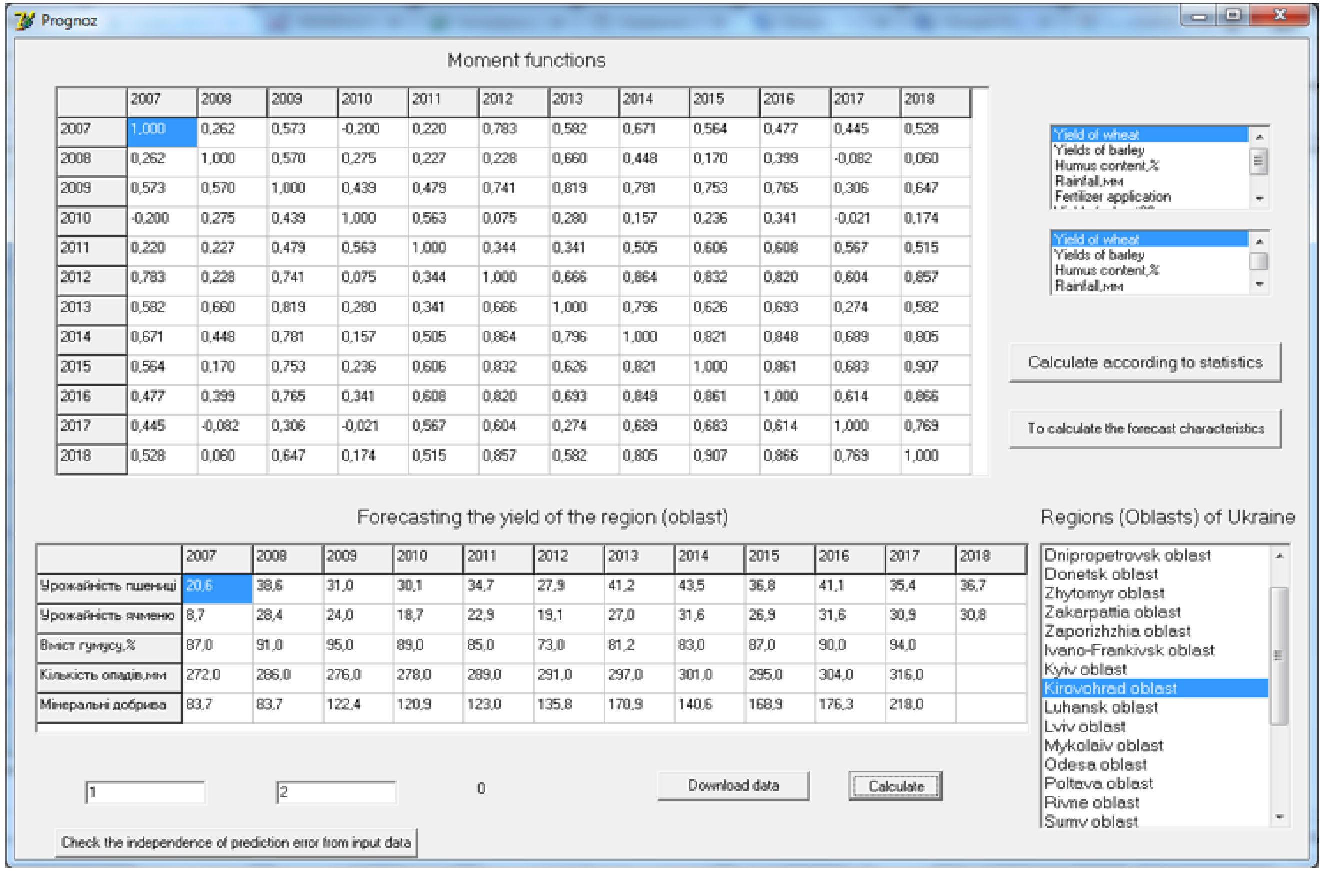
	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
2007	1,00	0,26	0,57	-0,19	0,21	0,78	0,58	0,67	0,56	0,47	0,44	0,52
2008	0,26	1,00	0,57	0,27	0,22	0,22	0,66	0,44	0,16	0,39	-0,08	0,05
2009	0,57	0,57	1,00	0,43	0,47	0,74	0,81	0,78	0,75	0,76	0,30	0,64
2010	-0,19	0,27	0,43	1,00	0,56	0,07	0,28	0,15	0,23	0,34	-0,02	0,17
2011	0,21	0,22	0,47	0,56	1,00	0,34	0,34	0,50	0,60	0,60	0,56	0,51
2012	0,78	0,22	0,74	0,07	0,34	1,00	0,66	0,86	0,83	0,82	0,60	0,85
2013	0,58	0,66	0,81	0,28	0,34	0,66	1,00	0,79	0,62	0,69	0,27	0,58
2014	0,67	0,44	0,78	0,15	0,50	0,86	0,79	1,00	0,82	0,84	0,68	0,80
2015	0,56	0,16	0,75	0,23	0,60	0,83	0,62	0,82	1,00	0,86	0,68	0,90
2016	0,47	0,39	0,76	0,34	0,60	0,82	0,69	0,84	0,86	1,00	0,61	0,86
2017	0,44	-0,08	0,30	-0,02	0,56	0,60	0,27	0,68	0,68	0,61	1,00	0,76
2018	0,52	0,05	0,64	0,17	0,51	0,85	0,58	0,80	0,90	0,86	0,76	1,00

$M\left[\overset{o}{X}_h(\nu)\overset{o}{X}_h(12)\right],\ \nu=\overline{1,11},\ h=\overline{1,2}$ are determined on the basis of determinate models:

$$M\left[\overset{o}{X}_1(\nu)\overset{o}{X}_1(12)\right]=0,614M\left[\overset{o}{X}_1(\nu)\overset{o}{X}_1(11)\right]+0,154M\left[\overset{o}{X}_1(\nu)\overset{o}{X}_1(10)\right]+$$
$$+0,041M\left[\overset{o}{X}_1(\nu)\overset{o}{X}_1(9)\right]-0,282M\left[\overset{o}{X}_1(\nu)\overset{o}{X}_1(8)\right],\ \nu=\overline{1,11}, \quad (8)$$

$$M\left[\overset{o}{X}_2(\nu)\overset{o}{X}_2(12)\right]=0,691M\left[\overset{o}{X}_2(\nu)\overset{o}{X}_2(11)\right]+0,088M\left[\overset{o}{X}_2(\nu)\overset{o}{X}_2(10)\right]-$$
$$-0,021M\left[\overset{o}{X}_2(\nu)\overset{o}{X}_2(9)\right]-0,012M\left[\overset{o}{X}_2(\nu)\overset{o}{X}_2(8)\right],\ \nu=\overline{1,11}, \quad (9)$$

Parameters of equations (8)-(9) meet the minimum of the average error of forecasting (the relative error doesn't exceed 1%) of the values of correlation functions and are obtained based on the processing of data for 5 years 2009-2017 using instrument "Search for solution" of Microsoft Excel table processor.



	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
2007	1,000	0,262	0,573	-0,200	0,220	0,783	0,582	0,671	0,564	0,477	0,445	0,528
2008	0,262	1,000	0,570	0,275	0,227	0,228	0,660	0,448	0,170	0,399	-0,082	0,060
2009	0,573	0,570	1,000	0,439	0,479	0,741	0,819	0,781	0,753	0,765	0,306	0,647
2010	-0,200	0,275	0,439	1,000	0,563	0,075	0,280	0,157	0,236	0,341	-0,021	0,174
2011	0,220	0,227	0,479	0,563	1,000	0,344	0,341	0,505	0,606	0,608	0,567	0,515
2012	0,783	0,228	0,741	0,075	0,344	1,000	0,666	0,864	0,832	0,820	0,604	0,857
2013	0,582	0,660	0,819	0,280	0,341	0,666	1,000	0,796	0,626	0,693	0,274	0,582
2014	0,671	0,448	0,781	0,157	0,505	0,864	0,796	1,000	0,821	0,848	0,689	0,805
2015	0,564	0,170	0,753	0,236	0,606	0,832	0,626	0,821	1,000	0,861	0,683	0,907
2016	0,477	0,399	0,765	0,341	0,608	0,820	0,693	0,848	0,861	1,000	0,614	0,866
2017	0,445	-0,082	0,306	-0,021	0,567	0,604	0,274	0,689	0,683	0,614	1,000	0,769
2018	0,528	0,060	0,647	0,174	0,515	0,857	0,582	0,805	0,907	0,866	0,769	1,000

	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Урожайність пшениці	20,6	38,6	31,0	30,1	34,7	27,9	41,2	43,5	36,8	41,1	35,4	36,7
Урожайність ячменю	8,7	28,4	24,0	18,7	22,9	19,1	27,0	31,6	26,9	31,6	30,9	30,8
Вміст гумусу,%	87,0	91,0	95,0	89,0	85,0	73,0	81,2	83,0	87,0	90,0	94,0	
Кількість опадів,мм	272,0	286,0	276,0	278,0	289,0	291,0	297,0	301,0	295,0	304,0	316,0	
Мінеральні добрива	83,7	83,7	122,4	120,9	123,0	135,8	170,9	140,6	168,9	176,3	218,0	

Fig. 3. Program interface for forecasting crop yield figures across regions of Ukraine

Table 2. Autocorrelative function of the component $X_2(i),\ i=\overline{1,12}$

	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
2007	1,00	0,11	0,59	0,52	0,15	0,84	0,68	0,72	0,76	0,61	0,67	0,74
2008	0,11	1,00	0,53	0,31	0,45	0,31	0,37	0,41	0,33	0,38	-0,08	0,22
2009	0,59	0,53	1,00	0,82	0,56	0,66	0,80	0,72	0,82	0,76	0,36	0,72
2010	0,52	0,31	0,82	1,00	0,49	0,64	0,79	0,67	0,78	0,73	0,40	0,64
2011	0,15	0,45	0,56	0,49	1,00	0,40	0,52	0,35	0,44	0,42	0,22	0,36
2012	0,84	0,31	0,66	0,64	0,40	1,00	0,83	0,91	0,87	0,80	0,69	0,77
2013	0,68	0,37	0,80	0,79	0,52	0,83	1,00	0,84	0,83	0,83	0,65	0,79
2014	0,72	0,41	0,72	0,67	0,35	0,91	0,84	1,00	0,91	0,92	0,69	0,80
2015	0,76	0,33	0,82	0,78	0,44	0,87	0,83	0,91	1,00	0,93	0,73	0,92
2016	0,61	0,38	0,76	0,73	0,42	0,80	0,83	0,92	0,93	1,00	0,69	0,88
2017	0,67	-0,08	0,36	0,40	0,22	0,69	0,65	0,69	0,73	0,69	1,00	0,82
2018	0,74	0,22	0,72	0,64	0,36	0,77	0,79	0,80	0,92	0,88	0,82	1,00

In Table 3, Table 4 coordinate functions $\beta_{h1}^{(h,1)}(\nu,i),\ \nu,i=\overline{1,12},\ h=\overline{1,2},$ corresponding to autocorrelated functions $M\left[\overset{o}{X}_h(\nu)\overset{o}{X}_h(i)\right],\ i=\overline{1,12},\ h=\overline{1,2}$ and determining the degree of influence of past values of wheat and barley productivity on future values of these figures are presented.

Consolidated results of quality of solving of the forecasting problem of cereal crop harvest across all regions of Ukraine are presented in Table 5.

Thus, results of the experiment show (Table 5) that application of non-linear relations in a predictive model allows significantly increase the quality of the forecasting of cereal crop harvest. The accuracy of determination of estimations of future values of crop yield parameters is 3-5 times higher as compared with a linear model.

If necessary a predictive model used for practical purposes can be easily modified by changing the settings of a software product (Fig. 3) and entering additional statistic data into Microsoft Excel file. For example, to increase the quality of solving of a forecasting problem, number I of discretization points t_i, N order of non-linear relations of an investigated random sequence, number H of components to take fuller account of the conditions of cereal crop production can be increased.

Table 3. Coordinate function $\beta_{11}^{(1,1)}(\nu,i)$ $\nu,i=\overline{1,12}$.

	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
2007	1,00	-0,26	0,43	0,35	-0,30	0,69	0,52	0,65	1,05	0,83	1,26	1,47
2008	0	1,00	0,90	0,37	0,94	-0,13	0,28	0,30	0,63	0,57	-0,89	0,67
2009	0	0	1,00	0,45	2,91	0,39	0,94	0,03	0,45	0,88	-0,02	0,80
2010	0	0	0	1,00	-1,77	0,12	0,06	0,03	-1,22	-0,71	-1,72	-3,58
2011	0	0	0	0	1,00	0,18	0,06	0,02	0,26	0,35	0,26	0,23
2012	0	0	0	0	0	1,00	0,22	0,80	1,01	0,22	4,60	2,40
2013	0	0	0	0	0	0	1,00	0,59	1,02	1,32	1,63	1,92
2014	0	0	0	0	0	0	0	1,00	2,29	0,79	3,45	3,53
2015	0	0	0	0	0	0	0	0	1,00	0,85	0,84	0,72
2016	0	0	0	0	0	0	0	0	0	1,00	0,30	0,03
2017	0	0	0	0	0	0	0	0	0	0	1,00	-1,02
2018	0	0	0	0	0	0	0	0	0	0	0	1,00

Table 4. Coordinate function $\beta_{21}^{(2,1)}(\nu,i)$ $\nu,i=\overline{1,12}$.

	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
2007	1,00	0,28	0,30	0,61	0,24	-0,03	-0,04	0,34	0,68	0,23	-0,11	0,10
2008	0	1,00	-0,25	-0,21	0,17	0,40	0,12	-0,85	-0,21	-0,93	-0,42	-0,63
2009	0	0	1,00	0,27	0,62	-0,21	0,63	0,62	0,49	0,69	1,17	0,28
2010	0	0	0	1,00	1,73	0,64	1,45	1,59	1,52	1,99	1,76	1,91
2011	0	0	0	0	1,00	1,82	1,02	1,12	0,21	-0,46	0,35	1,38
2012	0	0	0	0	0	1,00	1,84	1,07	-0,39	0,85	1,82	1,25
2013	0	0	0	0	0	0	1,00	-0,07	-0,38	-0,10	0,77	0,54
2014	0	0	0	0	0	0	0	1,00	0,47	0,61	0,00	0,29
2015	0	0	0	0	0	0	0	0	1,00	0,47	-0,22	0,30
2016	0	0	0	0	0	0	0	0	0	1,00	0,02	0,31
2017	0	0	0	0	0	0	0	0	0	0	1,00	-1,70
2018	0	0	0	0	0	0	0	0	0	0	0	1,00

Table 5. Relative errors of prognostication of cereal crop harvest.

Order of stochastic relations,	2	3	4
Relative error for wheat	6,9 %	3,2 %	1,5 %
Relative error for barley	7,1 %	3,3 %	1,6 %

5 Conclusion

Method of solving of an important economic problem of forecasting of cereal crop harvest is offered. A forecasting method, as well as an underlying canonical model, doesn't impose any limitations on the properties of a random sequence of change of crop yield figures (linearity, stationarity, linearity, Markov property, monotony, etc.). Taking into full account stochastic peculiarities of crop yield figures and conditions of cereal crop production allows to achieve maximum quality of solving of a forecasting problem. Results of a numerical experiment confirmed the high-accuracy characteristics of a predictive model for solving the problem of forecasting of crop yield figures for the regions of Ukraine. The model can also be used to improve the efficiency of the functioning of agricultural business enterprises. However, for microeconomic forecasting, it is necessary to modify a mathematical model taking into account the peculiarities of the economic activities of an agricultural enterprise (the composition and characteristics of the soil, weather conditions in the periods of climbing of cereals and ear crops, features of growing grain crops, taking into account the geographical location of an enterprise, etc. should be used as the parameters of a forecast model).

Application of an offered method of crop yield forecasting will allow to increase efficiency of the realization of Ukraine's Food Program and also to adjust the strategy of the use of cereal crops for production of alternative fuels [22,23]. For further studies, it is expedient to consider the possibility to use intellectual technologies [19,24] (a) for solving of the problems of crop yield forecasting and (b) for comparative analysis of the results of forecasting obtained on the basis of the approach offered in this work.

References

1. The main site of the agricultural business. www.latifundist.com
2. Gallego H., Kravchenko A.N., Kussul N.N., Skakun S.V, Shelestov A. Yu, Grypych Yu.V.: Efficiency Assessment of Different Approaches to Crop Classification Based on Satellite and Ground Observations. J. Problems of control and information sciences, №3, pp. 123-134 (2012)
3. Vijendra Kumar Boken: Forecasting Spring Wheat Yield Using Time Series Analysis. Agronomy Journal, Vol. 92, № 6, pp. 1047-1053 (2000)
4. Methodology of the MARS Crop Yield Forecasting System. Eur Rep 21291 EN/1-4 (2008)
5. Savin I., Negre T. (ed.): Agro-meteorological monitoring in Russia and Central Asian Countries, OPOCE EUR 22210EN, Ispra (Italy), 214 p. (2006)

6. Polovyi A.M.: Modelling of hydrometeorological regime and productivity of agroecosystems. Kyiv: KNT, 342 p. (2007)
7. Kogan F., Salazar L., Roytman L.: Forecasting crop production using satellite-based vegetation health indices in Kansas, USA. International Journal of Remote Sensing. 33, N 9. pp. 2798–2814 (2012)
8. Kacprzyk J.: Multistage Fuzzy Control: A Prescriptive Approach. John Wiley & Sons, Inc., New York, NY, USA (1997)
9. Shebanin V., Atamanyuk I., Kondratenko Y., Volosyuk Y.: Application of fuzzy predicates and quantifiers by matrix presentation in informational resources modeling. Proceedings of Conference "MEMSTECH 2016", 22-24 April 2016, Lviv-Poljana, pp. 146-149 (2016)
10. Tsay R.S.: Nonlinear time series models: testing and applications. Course in Time Series Analysis. New York: Wiley, 680 p. (2001)
11. Pugachev V.S.: The Theory of Random Functions and its Application. Fitmatgiz, Moscow, 720 p. (1962)
12. Box G.E.P., Jenkins G.M.: Time–series Analysis, Forecasting and Control. Holden–Day, San Francisco, 575 p. (1970)
13. Lozynsky O., Shcherbovskykh S.: Consideration effect of repair strategy on availability for non-redundant system using extended Markov model. Proceeding of Conference CADSM 2011, Polyana, Svalyava, Ukraine, pp. 2-22 (2011)
14. Wiener N.: Extrapolation, Interpolation, and Smoothing of Stationary Time Series. MIT Press, New-York, 163 p. (1970)
15. Kudritsky V.D.: Filtering, Extrapolation and Recognition Realizations of Random Functions. FADA Ltd., Kyiv, 176 p. (2001)
16. Atamanyuk I.P.: Algorithm of extrapolation of a nonlinear random process on the basis of its canonical decomposition. J. Kibernetika i Sistemnyj Analiz, №2, pp. 131-138 (2005)
17. Atamanyuk I.P.: Optimal Polynomial Extrapolation of Realization of a Random Process with a Filtration of Measurement Errors. Journal of Automation and Information Sciences. Volume 41, Issue 8, Begell House, USA, pp. 38-48 (2009)
18. Atamanyuk I.P., Kondratenko Y.P., Sirenko N.N.: Forecasting Economic Indices of Agricultural Enterprises Based on Vector Polynomial Canonical Expansion of Random Sequences. Proceedings of the 12th International Conference ICTERI'2016, CEUR-WS, Vol-1614, pp. 458-468 (2016)
19. Atamanyuk I., Kondratenko Y.: Calculation Method for a Computer's Diagnostics of Cardiovascular Diseases Based on Canonical Decompositions of Random Sequences. Proceedings of the 11th International Conference ICTERI-2015, Batsakis, S., et al. (Eds), CEUR-WS, Vol. 1356, Lviv, Ukraine, May 14-16, pp.108-120 (2015)
20. Atamanyuk I.P.: Polynomial algorithm of optimal extrapolation of stochastic system parameters. J. Upravlyayushchie Sistemy i Mashiny, №1, pp. 16-19 (2002)
21. State Statistics Service of Ukraine. www.ukrstat.gov.ua
22. Goncharuk A.G., Havrysh V.I., Nitsenko V.S.: National features for alternative motor fuels market. Int. J. of Energy Technology and Policy, Vol.14, No.2/3, pp. 226-249 (2018)
23. Kondratenko Y.P., Kozlov O.V., Kondratenko, G.V., Atamanyuk, I.P.: Mathematical Model and Parametrical Identification of Ecopyrogenesis Plant Based on Soft Computing Techniques, Complex Systems: Solutions and Challenges in Economics, Management and Engineering, Book Series: Studies in Systems, Decision and Control, Vol. 125, Berlin, pp. 201-233 (2018) DOI: https://doi.org/10.1007/978-3-319-69989-9_13
24. Trunov A.N.: An Adequacy Criterion in Evaluating the Effectiveness of a Model Design Process. Eastern-European Journal of Enterprise Technologies 1(4(73)), 36-41 (2015)