LATVIAN JOURNAL of PHYSICS and TECHNICAL SCIENCES

ISSN 0868 - 8257

6

(Vol. 59)

2022

CONTENTS

LATVIAN JOURNAL of PHYSICS and TECHNICAL SCIENCES

LATVIJAS FIZIKAS un TEHNISKO ZINĀTŅU ŽURNĀLS

ЛАТВИЙСКИЙ ФИЗИКО-ТЕХНИЧЕСКИЙ ЖУРНАЛ

Published six times a year since February 1964
Iznāk sešas reizes gadā kopš 1964. gada februāra
Выходит шесть раз в год с февраля 1964 года

6 (Vol. 59) • **2022**

RĪGA

INDEXED (PUBLISHED) IN

www.scopus.com
www.sciendo.com
EBSCO (Academic Search Complete, www.epnet.com), INSPEC (www.iee.org.com).
VINITI (www.viniti.ru), Begell House Inc/ (EDC, www.edata-center.com).

Issuers: Institute of Physical Energetics,
Institute of Solid State Physics, University of Latvia
Registration Certificate Number: 000700221

Editorial Contacts:
14 Dzerbenes Street, Riga, LV - 1006
Ph.: + 371 67551732
E-mail: leo@lza.lv
www.fei-web.lv

LATVIAN JOURNAL OF PHYSICS
AND TECHNICAL SCIENCES
2022, N 6

DOI: 10.2478/lpts-2022-0047

STATISTICAL MODELLING OF FACTORS INFLUENCING THE AGRICULTURAL LAND MARKET IN UKRAINE

Y. Lazebnyk[1*], O. Korepanov[1], T. Chala[1], G. Korepanov[1],
D. Chernenko[1], U. Plumite[2], M. Komlieva[1]

[1] V.N. Karazin Kharkiv National University,
Department of Statistics, Accounting and Auditing,
pl. Svobody, 4, 61022, Kharkiv, UKRAINE
[6]Riga Technical University,
[2]Institute of Civil Engineering and Real Estate Economics,
6-210 Kalnciema Str., LV-1048, Riga, LATVIA
*e-mail address: y.a.lazebnyk@karazin.ua

To make informed decisions, modern society, like modern business, must operate with adequate information about many complex interrelated aspects of its activities. Land use is only one of such aspects. Agricultural lands are of particular importance today. Land data are needed to analyse environmental processes, as well as end unsystematic, uncontrolled use of agricultural land, environmental degradation, destruction of important wetlands, loss of fish diversity and destruction of wildlife habitats. To solve these problems, it is advisable to conduct regular analysis and evaluation of land resources, justification and analysis of factors influencing the agricultural land market in the country. The tasks were solved using multidimensional statistical methods, in particular, factor analysis, which helped to get rid of subjectivity in the choice of factors influencing the market under study. Based on the model of influence of factors on the national market of agricultural land in Ukraine, six main factors were identified, the most influential of which were the components of agricultural development (first factor) and, to a lesser extent, the component of vegetable production and yield (fourth factor). The interrelationships and the degree of influence of the selected factors on the main indicators of the agricultural land market, namely, on the area of purchase and sale plots, the price of purchase and sale plots, and the number of transactions were also analysed.

***Keywords*:** *Agricultural lands, analysis of the development, factor analysis, land potential, land resources, method of principal components.*

1. INTRODUCTION

To make informed decisions, modern society, like modern business, must operate with adequate information about many complex interrelated aspects of its activities. Land use is only one of these aspects, but knowledge of land and soil cover continues to be increasingly important in terms of global food and energy security, environmental sustainability and economic growth.

Today agricultural lands are of particular importance. It is known that the development of infrastructure and the growing share of non-agricultural sectors of the economy significantly accelerate the process of withdrawal of land from agricultural use. This, in turn, gradually limits the area of arable land. As a result, the shortage of agricultural land is causing a decline in food production, which is exacerbated by rapid population growth in some parts of the world.

In the context of Ukraine, land data are needed to analyse environmental processes and end unsystematic, uncontrolled use of agricultural land, environmental degradation, destruction of important wetlands, loss of fish diversity and destruction of wildlife habitats. In addition, the importance of monitoring land relations is further determined by the opening of the market for agricultural land from 1 July 2021.

Given the above problems, there is a need for regular analysis and evaluation of land resources, justification factors influencing the market of agricultural land in Ukraine, thus proving the relevance of the study.

Issues related to theoretical developments and analysis of the development of the world market for agricultural land are considered in the works of such scientists as S. Balestri [1], M. Herold [2], S. Carter [2], Y. Zhang [6], H. Long [6], L. Ma [6], M. Maggioni [1], A. Menon [16], G. Robinson [9], J. Han [5], H. Lu [5], M. Vijayabaskar [3] and X. Zhang [5].

The research of B. McKay [7] and G. Oliveira [7] contributes to deepening of the understanding of land relations that go beyond the material value of land and are a way to protect national cultural identity.

Studies of C. Vaddhanaphuti [4], B. White [8], P. Vandergeest [8], E. Corbera [4], C. Park [8], L. Schoenberger [10], D. Hall [10] and C. Hunsberger [4] are devoted to the issue of harmonization of investments in agricultural lands with gender policy and climate change in some regions.

Despite the significant amount of modern research and development in this area, scientists are currently focusing on current issues of studying the current state and trends of the agricultural land market in Ukraine, which began operating in July 2021. In particular, it is important to consider the factors associated with the work of the object of study, which will assess the state of the land market in Ukraine and predict the main trends of its change.

2. THEORY METHODS AND METHODOLOGY

There is no doubt that the key to country's successful development, as well as its food security, is a rationally built and effective national agricultural policy, especially if the country has a powerful land potential.

In order to substantiate the state policy on the regulation of land relations, it is necessary to take into account the factors under which the state of the agricultural land market will be formed. This problem can be solved by using multidimensional statistical methods, in particular, factor analysis and principal component analysis (PCA), which help get rid of subjectivity in the choice of factors influencing the market under study.

Large datasets are increasingly common and are often difficult to interpret. PCA is a technique for reducing the dimensionality of such datasets, increasing interpretability but at the same time minimizing information loss. It does so by creating new uncorrelated variables that successively maximize variance. Finding such new variables, i.e., the principal components, reduces to solving an eigenvalue/eigenvector problem, and the new variables are defined by the dataset at hand, not *a priori*, hence, making PCA an adaptive data analysis technique. It is adaptive in another sense too, since variants of the technique have been developed that are tailored to various different data types and structures.

The principal components method is a method by which on the basis of really existing connections of features it is possible to reveal latent generalized characteristics of organisational structure and the mechanism of development of the phenomena and processes.

With the principal component method, the effect of multicollinearity can be eliminated by replacing the initial variables, which can be correlated with fewer independent components. These so-called hypothetical quantities cannot be directly estimated.

Tasks that are put forward before the component analysis and can be solved with its help are reduced to the following [17], [18]:

- reduction of initial variables as a result of allocation of principal components;
- identification of relationships and patterns, the so-called latent effects;
- evaluation of generalized indicators (interpretation and evaluation of the main components);
- predicting the development of a particular process, when constructing a regression equation, using the results of component analysis;
- typology and classification of objects of observation;
- use of component analysis in combination with other methods of analysis, i.e. data conversion for use in other models.

The model of component analysis can be represented as follows:

$$z_i = \sum_{i=1}^{m} a_{ik} G_k, \tag{1}$$

where

z_i – normalized value of the *i*-th variable with single variances;

a_{ik} – weight, factor load of the *k*-th component on the *i*-th variable; i = 1, 2,…, *m*; *k* = 1, 2,…, *p*;

m – number of variables;

p – number of principal components; $p<m$;

G_k – *k*-th principal component.

The algorithm of the principal components method consists of the following stages.

I. Formation of initial data matrix *X* size $n \times m$:

$$X = \begin{bmatrix} x_{11} x_{12} x_{13} \cdots\cdots x_{1m} \\ x_{21} x_{22} x_{23} \cdots\cdots x_{2m} \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ x_{n1} x_{n2} x_{n3} \cdots\cdots x_{nm} \end{bmatrix}, \tag{2}$$

where

x_{ij} – the value of the *i*-th variable for the *j*-th object (observation);

i = 1, 2,…, *m*, *m* – number of variables;

j = 1, 2,…, *n*, *n* – number of objects.

II. Calculation of the matrix of standardized values of variables, which is carried out according to the following formula:

$$z_{ij} = \frac{x_{ij} - \overline{x_i}}{\sigma_i}, \quad (3)$$

where
z_{ij} – standardized value of the i-th variable for the j-th object;
$\overline{x_i}$ – the average value of the i-th variable;
σ_i – standard deviation of the i-th variable.

III. Calculation of paired correlation coefficients matrix R with "1" on the main diagonal.

IV. In the principal components method, in contrast to factor analysis, it is believed that latent components should explain all variation. Therefore, there is no need to move from a correlation matrix R or a covariance matrix to a reduced covariance matrix or a reduced pair correlation matrix in which on the main diagonal instead of "1" there are communities (h_j^2).

Thus, the next step is to calculate the diagonal matrix of eigenvalues Λ size $p \times p$.

Inherent values λ_k indicate the contribution of the k-th component to the total variance of the initial data set.

V. Calculation of the orthogonal matrix of eigenvectors size $p \times p$:

$$U = \begin{pmatrix} u_{11} u_{12} u_{13} \cdots\cdots u_{1p} \\ u_{21} u_{22} u_{23} \cdots\cdots u_{2p} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ u_{p1} u_{pk2} u_{p3} \cdots\cdots u_{pp} \end{pmatrix} \quad (4)$$

VI. Calculation of factor loads matrix A, the elements of which are the weight of the components (factor loads):

$$A = V \Delta^{1/2} \quad (5)$$

VII. Before moving from the matrix A to the matrix of principal components values F, it may be necessary to find a simpler factor structure. The search for a simple factor structure is carried out using rotation procedures, as a result of which the values of some factor loads decrease and others increase. After that, a matrix of factor load after rotation is obtained.

VIII. Calculation of the matrix of principal component values F:

$$F = Z V \Delta^{-1/2}, \quad (6)$$

$$F = \begin{bmatrix} f_{11} f_{12} f_{13} \cdots\cdots f_{1k} \\ f_{21} f_{22} f_{23} \cdots\cdots f_{2k} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ f_{n1} f_{n2} f_{n3} \cdots\cdots f_{nk} \end{bmatrix}. \quad (7)$$

The main component method has certain properties that distinguish it from factor analysis. First, principal components are statistically independent. Second, the selected main components are ranked according to the level of their contribution to the total variance of the initial variables, i.e, the first main component has the maximum variance, the second – the largest variance among the remaining components, and so on until the full distribution of variance between the components.

Component analysis is a linear additive method. When using it, there is no need to hypothesize about the linearity of the model, the number of components and their correlation.

When using the principal components method, it is not necessary to make any assumptions about the variables, moreover – they can even be random variables.

With its help it is possible to completely decompose the variance of the initial variables, i.e., to fully explain it using latent components, which are generalized variables.

Provided that the weight of the components is determined and more than one of them – it is not always possible to unambiguously and adequately interpret the selected components. This is due to the fact that for the same components it is possible to obtain equivalent weights (loads) by their orthogonal transformation.

Transformations are performed in order to, if possible, find a simple factor structure. This is when the factor loads of traits that affect an unknown phenomenon have high values for one component and insignificant values for others. It should be noted that the element a_{ik} of the factor load matrix *A* indicates the relationship between the *i*-th initial variable and the *k*-th main component and is within the following limits: $-1 \leq a_{ik} \leq +1$ [17], [18]. The component for which the variables have high loads must be named, i.e., interpreted. In the case where we have more than one factor, they cannot always be interpreted unambiguously. In this regard, the factor structure can be changed using various procedures of orthogonal or oblique rotation, in the process of which the values of some factor loads increase and others decrease. The varimax procedure is most commonly used, which maximizes the variation of the factor load squares for each component by increasing the large and decreasing the factor load values. The factors that have the greatest load are given the appropriate name.

The following methods are used to select components whose solutions are easy to interpret.

1. Methods based on the rotation of factors (components) [17], [18]:

1. graphic method of rotation;
2. analytical methods:
   - varimax;
   - quartermax;
   - equimax;
   - biquartimax;

2. Methods that do not involve the rotation of components (factors).

These methods are based on the fact that before selection, the components must be specified as their number and characteristics, which must have zero weight on the components.

When deciding on the choice of the required number of components, it is advisable to use the Kaiser test or the Cattell method (the criterion of "rock collapse"). According to Kaiser's criterion, only those factors whose numbers are greater than one are left. According to Cattell's method, it is necessary to graphically display the eigenvalues of the correlation matrix in descending order. The selection of factors ends with the factor after which the inherent numbers of the correlation matrix do not decrease rapidly.

If, after taking into account, for example, 75% of the variance, the next component explains less than a given percentage of variance, it is excluded because it has too little contribution to the total variance, and therefore this component is not important.

Due to the fact that principal components are orthogonal to each other, the total variance of the *i*-th variable is equal to one. In this regard, the contribution of the *k*-th component to the total variance can be determined as follows [17], [18]:

$$\frac{\lambda_k}{m} 100\%. \tag{7}$$

The total contribution of the selected principal components or, otherwise, the completeness of the factorization is determined as follows [17], [18]:

$$\frac{\sum_{k=1}^{p} \lambda_k}{m} 100\%. \tag{8}$$

As for the significance of the weight of the components (factor loads), in socio-economic studies they can be considered significant if they are not less than a predetermined value.

For economic interpretation, only those factor loads are used, the values of which are greater 0.7 [18].

Using the theoretical provisions of factor analysis, namely, the principal component method, the main factors influencing the market of agricultural land were identified: reduction of the space of selected features; selection of a small number of uncorrelated components that store all the information on the causal mechanism of the phenomenon; interpretation and evaluation of principal components; application of the principal components method in combination with other multidimensional methods of analysis, namely multiple regression analysis.

The model of the influence of factors on the national market of agricultural land in Ukraine was developed using regional indicators published by the State Statistics Service of Ukraine [11]–[14], which could affect researched market.

The variables in this model are defined as the average annual values of the following indicators by regions for 2015–2020:

- X1 – Present population, thousand people;
- X2 – Indices for agricultural production, %;
- X3 – Indices for crop production, %;
- X4 – Indices for livestock production, %;
- X5 – Share of regions in the agricultural production, %;
- X6 – Share of agricultural production by enterprises;
- X7 – Share of crop production, %;
- X8 – Labour productivity at enterprises that were engaged into agricultural activity (per 1 employed in agricultural production, at constant prices in 2016; thousand UAH);
- X9 – Share of regions in the crop production, %;
- X10 – Share of crop production by enterprises, %;
- X11 – Sown area for agricultural crops, thousand hectares;
- X12 – Cereals and leguminous crops output, thousand tons;
- X13 – Yields of cereals and leguminous crops, c of 1 ha;
- X14 – Sunflower production, thousand tons;
- X15 – Yields of sunflower, c of 1 ha;
- X16 – Potatoes production, thousand tons;
- X17 – Yields of potatoes, c of 1 ha;
- X18 – Vegetables crops production, thousand tons;
- X19 – Yields of vegetables crops, c of 1 ha;
- X20 – Fruit and berry crops production, thousand tons;
- X21 – Yields of fruit and berry crops, c of 1 ha;
- X22 – Share of regions in the livestock production, %;
- X23 – Share of livestock production by enterprises, %.

The average annual population in the regions of Ukraine in 2015–2020 was calculated according to the formula of the chronological average according to the current values of the indicator (at the beginning of 2015–2021) [11], [13].

Average annual indices of agricultural products, indices of crop and livestock products was calculated according to the geometric mean formula based on chain growth rates in 2015–2020 [12], [14].

Average annual labour productivity in enterprises engaged in agricultural activities in 2015–2020 was calculated as the arithmetic mean weighted by the number of the employed in agricultural production in the year [12], [14].

Average annual yields of cereals and legumes, sunflower, potatoes, vegetables, fruits and berries were calculated by the formula of the average harmonic, the numerator of which was the production of relevant crops [12], [14].

3. RESULTS AND DISCUSSION

Statistical model of factors influencing the market of agricultural land in Ukraine was built using the application package "STATISTICA" version 10, in particular the module "Factor Analysis".

As a result of the implementation of step-by-step procedures of factor analysis, the following results were obtained.

At the first stage, they were calculated values of eigenvalues λ_j. Principal components are those for which the Kaiser test λ_j>1 [17], [18]. Therefore, only those factors are taken into account, the inherent numbers of which are greater than one. There are six such factors (Table 1).

Table 1. Values of Eigenvalues and Contribution of Each of the Variances to the Total Variation of the Sign Set

Factor	Values of eigen-values	Proportion of total variance, %	Accumulated values of eigenvalues	Accumulated particles of total dispersion, %
1	7.11	30.89	7.11	30.89
2	5.71	24.83	12.82	55.72
3	3.06	13.31	15.88	69.03
4	2.05	8.90	17.93	77.94
5	1.42	6.17	19.35	84.11
6	1.15	4.99	20.49	89.10

Source: Built on own calculations according to data [11]–[14].

This is also confirmed by the graphical criterion of "rock collapse" [17], [18] (Fig. 1). The six selected main components explain 89.1 % of the total variation, which indicates a high degree of factorization.

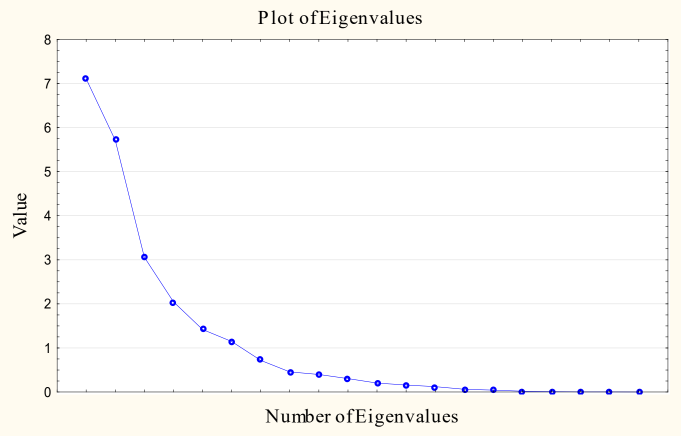

Fig. 1. Graphic display values of eigenvalues.
Source: Built by the authors according to their own calculations by the data [11]–[14].

Next, the factor loads were calculated, which have the meanings given in Table 2.

Table 2. Factor Loads before the Rotation of Factors*

Variables	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
X1	–0.110	0.425	0.416	0.080	–0.049	**0.711**
X2	0.047	–0.587	–0.615	0.357	0.040	0.132
X3	0.212	–0.434	**–0.715**	0.314	–0.026	0.233
X4	–0.355	–0.617	0.411	0.125	0.260	-0.300
X5	**–0.969**	–0.015	0.123	0.162	–0.005	–0.107
X6	**–0.849**	0.216	–0.285	–0.215	0.081	0.198
X7	–0.378	0.557	–0.573	–0.230	–0.337	–0.071
X8	–0.427	**–0.593**	–0.400	–0.052	0.009	0.280
X9	**–0.959**	0.137	–0.038	0.073	–0.107	–0.166
X10	**–0.797**	0.210	–0.400	–0.285	0.018	0.150
X11	**–0.777**	0.571	–0.034	0.073	–0.045	–0.124
X12	**–0.946**	0.115	–0.060	–0.091	–0.073	–0.169
X13	–0.395	**–0.834**	–0.089	–0.225	0.000	–0.144
X14	–0.621	**0.673**	–0.041	–0.011	0.000	–0.080
X15	–0.429	**–0.829**	0.033	–0.215	–0.073	–0.052
X16	–0.371	**–0.825**	0.009	0.033	0.031	0.087
X17	–0.048	**–0.812**	–0.165	–0.235	–0.340	–0.002
X18	–0.270	0.271	–0.096	**0.834**	0.019	–0.080
X19	0.084	0.038	–0.468	**0.712**	–0.258	–0.138
X20	–0.130	–0.296	0.606	0.290	–0.469	–0.052
X21	–0.137	–0.121	0.397	0.007	**–0.818**	0.179
X22	–0.641	–0.371	0.468	0.297	0.260	0.057
X23	**–0.701**	–0.129	0.261	0.206	0.251	0.390
λj	7.105	5.711	3.062	2.048	1.420	1.148
Proportion of total variance.%	30.9	24.8	13.3	8.9	6.2	5.0

Source: Authors' own calculations for [11]–[14].

* – Excluding the temporarily occupied territory of the Autonomous Republic of Crimea, the city of Sevastopol and a part of the temporarily occupied territories in the Donetsk and Luhansk regions.

Analysing Table 2, it is possible to determine that the obtained results are difficult to interpret and for a better understanding and correct interpretation of the results it is advisable to use the procedure of rotation of factors. "Varimax normalized" was chosen from the rotation procedures.

Inherent numbers and contribution of individual components to the total variance of the transformed factor loads are given in Table 3.

Table 3. Factor Loads after Rotation of Factors* (Varimax Normalized)

Variables	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
X1	0.092	–0.348	0.003	–0.117	0.215	**0.831**
X2	–0.180	0.679	–0.038	0.565	–0.239	–0.023
X3	–0.254	0.592	–0.278	0.578	–0.264	0.051
X4	0.013	0.313	**0.825**	–0.101	0.083	–0.232
X5	**0.834**	0.132	0.493	0.089	0.162	0.038
X6	**0.894**	0.199	–0.031	–0.091	–0.183	0.242
X7	**0.711**	–0.084	–0.649	0.091	–0.039	–0.104
X8	0.234	**0.819**	0.072	0.103	–0.111	0.147
X9	**0.935**	0.083	0.267	0.095	0.146	–0.045
X10	**0.886**	0.248	–0.161	–0.095	–0.196	0.156
X11	**0.914**	–0.320	0.061	0.091	0.035	0.049
X12	**0.935**	0.133	0.216	–0.046	0.090	–0.074
X13	0.117	**0.831**	0.349	–0.173	0.052	–0.268
X14	**0.812**	–0.418	–0.054	0.014	–0.039	0.083
X15	0.121	**0.814**	0.385	–0.209	0.176	–0.161
X16	0.026	**0.778**	0.461	0.004	0.097	0.004
X17	–0.156	**0.839**	–0.012	–0.106	0.274	–0.213
X18	0.272	–0.301	0.255	**0.788**	0.020	0.086
X19	–0.012	0.004	–0.164	**0.876**	0.060	–0.148
X20	–0.097	0.031	0.418	0.059	**0.768**	0.012
X21	0.050	0.114	–0.076	–0.063	**0.919**	0.153
X22	0.292	0.196	**0.862**	0.018	0.126	0.208
X23	0.452	0.180	0.570	0.020	0.003	0.522
λj	6.616	4.965	3.402	2.222	1.897	1.392
Proportion of total variance.%	28.8	21.6	14.8	9.7	8.2	6.1

Source: Authors' own calculations for [11]–[14].

* – Excluding the temporarily occupied territory of the Autonomous Republic of Crimea, the city of Sevastopol and a part of the temporarily occupied territories in the Donetsk and Luhansk regions.

As can be seen from Table 3, the contribution of the first component in the total variance of the data set is 28.8 %, the second – 21.6 %, the third – 14.8 %, the fourth – 9.7 %, the fifth – 8.2 %, the sixth – 6.1 %. Together, these components account for almost 89.2 % of the total variation, which indicates a high level of factorization. Note that the first component is closely related to the variables: X5, X6, X7, X9, X10, X11, X12, X14. Thus, the first component can be interpreted as a generalized factor in the development of agriculture with an emphasis on crop production. Another component is closely related to the variables: X8, X13, X15, X16. It is interpreted as a factor in the intensive development of agriculture. The third component is closely related to the following variables: X4, X22. It can be interpreted as a factor in rural development livestock. The fourth component is closely related to the variables: X18, X19. It is interpreted as a factor in the production and yield of vegetable crops. The fifth component is closely related to the variables: X20, X21. It is interpreted as a factor in the production and yield of fruit and berry crops. The sixth component is closely related to the variable: X1. It is interpreted as a factor in the size of the existing population. The values of principal components for the regions of Ukraine are given in Table 4.

Table 4. Values of Principal Components for the Regions of Ukraine*

Region (oblast)	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
Vinnytska	1.002	0.498	2.413	0.240	1.179	–0.352
Volynska	–1.145	0.495	0.192	0.231	–0.870	0.392
Dnipropetrovska	0.698	–1.023	1.019	1.034	0.061	1.807
Donetska	–0.462	–1.084	–0.400	–1.306	0.606	2.867
Zhytomyrska	–0.302	1.128	–0.530	0.767	–0.050	–0.335
Zakarpatska	–2.363	–1.189	0.728	–0.003	0.521	–1.470
Zaporizhska	0.339	–1.073	–0.654	0.324	–0.808	0.301
Ivano–Frankivska	–1.605	0.291	0.591	–0.655	–1.195	0.106
Kyivska	0.331	0.184	1.616	–0.219	–0.531	0.390
Kirovohradska	0.946	–0.945	–0.070	–1.332	–1.075	–1.467
Luhanska	–0.211	0.019	–2.627	–0.051	0.074	1.042
Lvivska	–1.030	0.885	0.515	0.657	–0.165	1.185
Mykolayivska	0.519	–1.465	–0.630	0.795	–0.750	–1.162
Odeska	0.779	–1.856	–0.558	–1.883	0.881	–0.572
Poltavska	1.319	0.520	–0.324	0.206	2.085	–0.450
Rivnenska	–1.027	0.865	–0.496	0.021	0.454	–0.084
Sumska	0.758	1.353	–0.492	–0.615	–1.476	–0.517
Ternopilska	–0.047	1.101	–0.493	–0.037	0.331	–0.406
Kharkivska	1.035	–0.104	–0.082	0.637	0.464	0.462
Khersonska	0.261	–1.121	–0.213	3.184	–0.523	–0.435
Khmelnytska	0.543	1.374	–0.399	–0.099	1.745	–0.203
Cherkaska	0.566	0.043	1.565	–1.139	–0.928	0.268
Chernivetska	–1.696	–0.233	–0.071	–0.337	1.412	–1.069
Chernihivska	0.793	1.337	–0.601	–0.421	–1.443	–0.297

Source: Authors' own calculations for [11]–[14].

* – Excluding the temporarily occupied territory of the Autonomous Republic of Crimea, the city of Sevastopol and a part of the temporarily occupied territories in the Donetsk and Luhansk regions.

Estimates of principal components are used to measure the relationship and study the impact of selected factors on the main indicators of the agricultural land market, namely, the area of plots for sale from 01.07.2021 to 16.01.2022, ha (Y_1), the price of plots of land for sale for the period from 01.07.2021 to 16.01.2022, UAH per hectare (Y_2), and the number of transactions for the period from 01.07.2021 to 16.01.2022 (Y_3).

Describing the matrix of correlation coefficients for the three dependent (area of sales, price, and number of transactions) and independent variables (factors), it should be noted the high level of correlation (positive relationship) between the dependent variables of the number and area of sales is a logical explanation of the mechanism of functioning of the market of agricultural lands and the specifics of reporting data on it (Table 5).

Table 5. Correlation Coefficients between the Main Indicators of the Agricultural Land Market and the Selected Factors

	Y_1	Y_2	Y_3
Y_1	1.00		
Y_2	–0.22	1.00	
Y_3	0.88	–0.13	1.00
Factor 1	0.62	–0.17	0.70
Factor 2	–0.25	0.24	0.10
Factor 3	–0.05	0.39	0.05
Factor 4	0.41	–0.05	0.28
Factor 5	0.04	0.00	0.06
Factor 6	0.08	0.22	–0.04

Source: Authors' own calculations for [11]–[14].

An increase in the number of transactions leads to an increase in the total area of sales plots, but this relationship may not be fully functional, because the area of individual transactions is different and can vary significantly. In addition, despite the low in absolute terms statistically insignificant correlation between the area of purchase and sale and their price, the negative direction of this relationship is clear and is interpreted in the laws of supply and demand – at a high price fewer land transactions and increase in the total area of land sold (increase in supply) lead to a gradual decline in market prices for agricultural land.

Regarding the relationships between independent and dependent variables, there are high enough correlation coefficients for the first factor (first component) in the equations with the area and number of sales transactions. Since the first component is interpreted as a generalized factor in the development of crop production, this correlation characterises the main targeted use of agricultural land in Ukraine and confirms the market activity of those companies that are primarily related to crop production.

Using regression analysis to study the impact of the six main components (factors) of agriculture identified above, three main models were built to characterise each of the three dependent variables separately.

The multiple regression equation looks like this:

$$Y_i = a_0 + a_1G_1 – a_2G_2 – a_3G_3 + a_4G_4 + a_5G_5 + a_6G_6, \quad (9)$$

where

Y_i – theoretical values of the chosen ones;

a_0 – initial regression parameter;

a_j – regression parameter of the factor G_j, $i = 1, …, 6$;

G_j – factor selected by the principal components method, $i = 1, …, 6$.

The first regression model considers the effect of six factors on the area of purchase and sale. The significance of the coefficients of the model was estimated in Excel using the statistical procedure Regression in the Data Analysis package. The results of the calculations are presented in Table 6.

Table 6. Characteristics of the Regression Model Areas of Plots of Sale of Agricultural Land in Ukraine (y_1)

Regression Statistics						
Multiple R	0.789					
R Square	0.624					
Adjusted R Square	0.491					
Standard Error	3183.118					
Observations	24					
ANOVA (Analysis of variance)						
	df	***SS***	***MS***	***F***	***Significance F***	
Regression	6	285839766.3	47639961.06	4.702	0.00539	
Residual	17	172248090.4	10132240.61			
Total	23	458087856.7				
	Coefficients	***Standard Error***	***t Statistics***	***P-value***	***Lower 95%***	***Upper 95%***
Y-section	4398.97	649.751	6.770	0.0000	3028.12	5769.83
Factor 1	2765.24	663.726	4.166	0.0006	1364.90	4165.58
Factor 2	–1104.27	663.726	–1.664	0.1145	–2504.61	296.07
Factor 3	–217.29	663.726	–0.327	0.7474	–1617.63	1183.05
Factor 4	1827.38	663.726	2.753	0.0136	427,04	3227.72
Factor 5	188.23	663.726	0.284	0.7801	–1212.11	1588.57
Factor 6	374.01	663.726	0.563	0.5805	–1026.33	1774.35

Source: Authors' own calculations for [15], [11]–[14].

Checking the materiality of the connection is statistically formulated as a test of null hypotheses: H_0: $R^2 = 0$; H_0: $a_j = 0$. Hypothesis H_0 is rejected or accepted on the basis of statistical criteria, in particular the variance F-criterion, the statistical characteristic of which is calculated by the ratio of estimates of factor and residual variances [17], [18]. Critical values $F_{1-\alpha}(k_1, k_2)$, where α – the materiality level, $k_1 = m–1$, $k_2 = n–(m–1)$ – the number of freedom degrees of the numerator and denominator was determined using the EXCEL FINV function. In our case $F_{emp.} = 4.7 > F_{crit.}$ (1–α; k_1; k_2) = 2.77.

The strength of the connection can be inferred from the value of the coefficient of determination R^2. For the model obtained (Table 6):

- close to unity coefficient of determination $R^2 = 0.62$,
- the estimated value of F-statistics ($F_{emp.}$ = 4.7) is more than critical $F_{åmp.} = 4.7 > F_{crit.} = 2.7$, which indicates the high adequacy of the constructed model.

The analysis allows us to conclude that the constant a_0 and the coefficient a_1 and a_4 are significant, because the absolute values of their t-statistics are more than covered, because the hypothesis H_0 at the level of significance α deviates if the inequality $t_{emp.} > t_{crit.1-\alpha;\, n-m-1}$ holds, where $t_{emp.}$ is calculated and $t_{crit.\varepsilon; n-m-1}$ is determined from the table of theoretical values of the t-test at the level of significance α and ($n–m–1$) degrees of freedom. Theoretical significance of the t-test $t_{crit.}$ was obtained using the function EXCEL T.INV.2T ((1–α), $n–m–1$) = T.INV.2T (0,05, 24–6–1) = 2.11 (for two-way distribution).

Therefore, the equation of multiple

regression on the selected factors is as follows:

$$Y_1 = 4398.9+2765.2\ G_1–1104.2\ G_2–217.3\ G_3+1827.4\ G_4+188.2\ G_5+374\ G_6.$$

Considering the regression coefficients on the factors, it is seen that only the first (generalized factor of agricultural development with an emphasis on crop production) and the fourth (factor of production and yield of vegetable crops) factors affect the area of purchase and sale, because *t*-statistics of these coefficients are higher than the corresponding tabular value of 2.11 for 95 % of the confidence interval.

The significance of the other four factors has not been confirmed. Thus, we can talk about the real positive (due to the sign of the coefficient) impact of the first and fourth main components on the area of purchased/sold plots of agricultural land in Ukraine.

Analysing the second regression model (Table 7), which describes the impact of six components on the price of plots of sale, no statistically significant effect was found.

Table 7. Characteristics of the Regression Model Prices of Plots for Sale of Agricultural Land in Ukraine (y_2)

Regression Statistics						
Multiple R	0.540					
R Square	0.292					
Adjusted R Square	0.042					
Standard Error	13155.943					
Observation	24					
ANOVA (Analysis of variance)						
	df	***SS***	***MS***	***F***	***Significance F***	
Regression	6	1213657674	202276279.0	1.169	0.368	
Residual	17	2942340659	173078862.3			
Total	23	4155998333				
	Coefficients	***Standard Error***	***t Statistics***	***P-value***	***Lower 95 %***	***Upper 95 %***
Y-section	35604.25	2685.446	13.258	0.0000	29938.45	41270.05
Factor 1	–2319.53	2743.204	–0.846	0.4095	–8107.18	3468.13
Factor 2	3289.55	2743.204	1.199	0.2469	–2498.11	9077.20
Factor 3	5230.39	2743.204	1.907	0.0736	–557.26	11018.05
Factor 4	–738.84	2743.204	–0.269	0.7909	–6526.49	5048.82
Factor 5	41.56	2743.204	0.015	0.9881	–5746.09	5829.22
Factor 6	2943.09	2743.204	1.073	0.2983	–2844.56	8730.74

Source: Authors' own calculations for [15], [11]–[14].

The third regression model examines the effect of six factors on the number of sales transactions, and the results of the calculations are presented in Table 8.

Table 8. Characteristics of the Regression Model and the Number of Transactions of Purchase and Sale of Agricultural Land in Ukraine (y_3)

Regression statistics						
Multiple R	0.765					
R Square	0.585					
Adjusted R Square	0.439					
Standard Error	729.180					
Observation	24					
ANOVA (Analysis of variance)						
	df	*SS*	*MS*	*F*	*Significance F*	
Regression	6	12764336.55	2127389.43	4.001	0.011	
Residual	17	9038957.29	531703.37			
Total	23	21803293.83				
	Coefficients	*Standard Error*	*t Statistics*	*P-value*	*Lower 95%*	*Upper 95%*
Y-section	1210.08	148.843	8.130	0.000	896.056	1524.115
Factor 1	680.73	152.045	4.477	0.000	359.949	1001.521
Factor 2	97.51	152.045	0.641	0.530	–223.276	418.296
Factor 3	53.10	152.045	0.349	0.731	–267.683	373.889
Factor 4	272.92	152.045	1.795	0.090	–47.863	593.709
Factor 5	59.42	152.045	0.391	0.701	–261.365	380.207
Factor 6	–35.00	152.045	–0.230	0.821	–355.790	285.782

Source: Authors' own calculations for [15], [11]–[14].

The adequacy of such a model is slightly lower ($R^2 = 0.585$) than when determining the effects of components on the area of purchase and sale, but the actual value of the F-criterion is greater than the F tabular (4.0 > 2.77), so the model is statistically significant and in general it describes 58.5 % of the variation in the dependent variable. The multiple regression equation looks like this:

$$Y_3 = 1210.08 + 680.73G_1 + 97.50G_2 + 53.10G_3 + 272.92G_4 + 59.42G_5 – 35G_6.$$

Unlike the first model, the regression coefficient for the factor of production and yield of vegetable crops (component four) is no longer statistically significant at 95 % confidence interval, and only the influence of the factor of agricultural development with emphasis on crop production remains different from zero (value of t-statistics of 4.48 is higher than the corresponding table value of 2.11). Thus, the impact of the six studied components on the three main indicators of the market of purchase and sale of agricultural land in Ukraine is limited mainly by the effects of the agricultural development component, and the development and dynamics of land prices are formed entirely without the influence of all identified components.

4. CONCLUSIONS

The generalization of the theoretical foundations of the basic concepts and categories used in the study has allowed us to determine that relevant to the research topic is the category of agricultural land, which, in turn, is divided into agricultural and non-agricultural land. Both types of land are subject to the land market in Ukraine through the NMV mechanism, but the level of statistical reporting and availability of information on the latter is insufficient compared to the former, which requires improved approaches to collecting and disseminating relevant statistical information.

Factor analysis proves a statistically significant impact of two of the six most important selected agricultural components (agricultural development components (first factor) and to a lesser extent vegetable production and yield component (fourth factor)) on the number and area of agricultural land sales in Ukraine. None of the considered components, however, has a statistically significant impact on the development of prices for agricultural land.

REFERENCES

1. Balestri, S., & Maggioni, M. A. (2017). This Land is my Land! Large-Scale Land Acquisitions and Conflict Events in Sub-Saharan Africa. *2017 World Bank Conference on Land and Poverty.* Washington DC. 20–24 March 2017. Available at https://www.researchgate.net/publication/314871867_This_land_is_my_land_Large-Scale_Land_Acquisitions_and_conflict_events_in_Sub-Saharan_Africa
2. Carter, S., Manceur, A. M., Seppelt, R., Hermans-Neumann, K., Herold, M., & Verchot, L. (2017). Large-Scale Land Acquisitions and REDD+: A Synthesis of Conflicts and Opportunities. *Environmental Research Letters*, *12* (3). Available at https://www.cifor.org/publications/pdf_files/articles/ACarter1802.pdf
3. Cherniak, O. I., & Kudinenko, M. Ya. (2002). Analysis and Forecast of Ukraine's GDP Dynamics Using the Method SSA (in Ukrainian). *Ekonomika i prohnozuvannia*, *4*, 134–147. Institute of Economics and Forecasting. Kyiv [Online]. Available at https://econom.univ.kiev.ua/wp-content/uploads/science/articles/Chernyak_OI/Analiz_ta_prognoz_dynamiky_VVP_Urainy_za_dopomogou_metody_SSA.pdf
4. Corbera, E., Hunsberger, C., & Vaddhanaphuti, C. (2017). Climate Change Policies, Land Grabbing and Conflict: Perspectives from Southeast Asia. *Canadian Journal of Development Studies, 38* (3)*,* 297–304.
5. Han, J., Jiang, M., Zhang, X., & Lu, X. (2021). Knowledge Mapping Analysis of Transnational Agricultural Land Investment Research. *Land Use Policy*, *10* (12). Available at https://www.mdpi.com/2073-445X/10/12/1374/htm
6. Long, H., Zhang, Y., Ma, L., & Tu, S. (2021). Land Use Transitions: Progress, Challenges and Prospects. *Land Use Policy*, *10* (9). Available at https://webcache.googleusercontent.com/search?q=cache:jBzxLcxZSZkJ:https://www.mdpi.com/2073-445X/10/9/903/pdf+&cd=2&hl=ru&ct=clnk
7. McKay, B. M., Oliveira, G. d. L. T., & Liu, J. (2020). Authoritarianism, Populism, Nationalism and Resistance in the Agrarian South. *Canadian Journal of Development Studies*, 41, 347–362.
8. Park, C. M. Y., & White, B. (2017). Gender and Generation in Southeast Asian Agro-commodity Booms. *The Journal of Peasant Studies*, 1103–1110.

9. Robinson, G.M., & Carson, D.A. (2015). The Globalisation of Agriculture: Introducing the Handbook. *Annu Rev Resour Econ*., 1–28.
10. Schoenberger, L., Hall, D., & Vandergeest, P. (2017). What Happened when the Land Grab Came to Southeast Asia? *The Journal of Peasant Studies*, *44*, 697–725.
11. State Statistics Service of Ukraine. (2022). *Rehiony Ukrainy 2019. Statystychnyj zbirnyk [Regions of Ukraine 2019. Statistical yearbook]* (in Ukrainian), section 1, Derzhkomstat, Kyiv, Ukraine. Available at http://www.ukrstat.gov.ua/druk/publicat/kat_u/2019/zb/12 /zb_ru1ch2019.pdf
12. State Statistics Service of Ukraine. (2022). *Rehiony Ukrainy 2019. Statystychnyj zbirnyk [Regions of Ukraine 2019. Statistical yearbook]*, section 2, Derzhkomstat, Kyiv, Ukraine. Available at http://www.ukrstat.gov.ua/druk/publicat/kat_u/2019/zb/12/zb_ru2ch2019.pdf
13. State Statistics Service of Ukraine. (2022). *Rehiony Ukrainy 2020. Statystychnyj zbirnyk [Regions of Ukraine 2020. Statistical yearbook]* (in Ukrainian), section 1, Derzhkomstat, Kyiv, Ukraine. Available at https://ukrstat.org/uk/druk/publicat/kat_u/2021/zb/12/Regionu_20_pdf.zip
14. State Statistics Service of Ukraine. (2022). *Rehiony Ukrainy 2020. Statystychnyj zbirnyk [Regions of Ukraine 2020. Statistical yearbook]* (in Ukrainian), section 2, Derzhkomstat, Kyiv, Ukraine. Available at https://ukrstat.org/uk/druk/publicat/kat_u/2021/zb/12/Regionu_20_pdf.zip
15. The Official Agricultural and Political Site of Ukraine. (2022). *Land Market in Ukraine. Map of Land Agreements and Land Prices from July 1, 2021* (in Ukrainian). Available at https://agropolit.com/spetsproekty/892-rinok-zemli-v-ukrayini-mapa-zemelnih-ugod-i-tsin-na-zemlyu-z-1-lipnya-2021-roku?utm_source=kurkul&utm_medium=news
16. Vijayabaskar, M., & Menon, A. (2018). Dispossession by Neglect: Agricultural Land Sales in Southern Indian. *J. Agrar. Chang.*, 18, 571–587.
17. Yerina, A.M. (2001). *Statystychne modeliuvannia ta prohnozuvannia* [Statistical modelling and forecasting] (in Ukrainian): navch. posib. KNEU, Kyiv, Ukraine.
18. Yerina, A.M., & Yerin, D.L. (2014). *Statystychne modeliuvannia ta prohnozuvannia* [Statistical modelling and forecasting] (in Ukrainian): navch. posib. KNEU, Kyiv, Ukraine.