

Електронний журнал «Ефективна економіка» включено до переліку наукових фахових видань України з питань економіки (Категорія «Б», Наказ Міністерства освіти і науки України № 975 від 11.07.2019). Спеціальності – 051, 071, 072, 073, 075, 076, 292. Ефективна економіка. 2022. № 6.

DOI: 10.32702/2307-2105.2022.6.24

УДК: 330.4

*Л. Б. Прокопович,
к. е. н., доцент, доцент кафедри «Облік і економічний аналіз»,
Національний університет кораблебудування імені адмірала Макарова
ORCID ID: 0000-0003-2561-8862*

ОЦІНЮВАННЯ ОБСЯГУ АГРАРНОГО ВИРОБНИЦТВА

*L. Prokopovich,
PhD in Economics, Associate Professor,
Associate Professor of the Department of Accounting and Economic Analysis,
Admiral Makarov National University of Shipbuilding*

ASSESSMENT OF THE VOLUME OF AGRICULTURAL PRODUCTION

Стаття присвячена актуальній проблемі підвищення достовірності оцінювання рівня обсягу аграрного виробництва. У роботі був проаналізований характер існуючої залежності рівня обсягу аграрного виробництва України. Як фактори, що впливають на величину обсягу аграрного виробництва, розглядалися: зайняті в сільськогосподарському виробництві, вартість основних засобів, посівні площі сільськогосподарських культур. Була побудована система лінійних і нелінійних однофакторних та багатфакторних регресійних моделей. Проведений аналіз р-значень для внутрішніх параметрів моделей. Дослідження величин середньої помилки апроксимації побудованих моделей виявив їх перевищення граничного значення, що вказує на

неможливість використання отриманих моделей для прогнозування обсягу аграрного виробництва. Побудовані багатофакторні регресійні моделі із використанням методів машинного навчання з вчителем: k-найближчих сусідів і випадковий ліс. Модель на основі випадкового лісу має допустиму величину помилки апроксимації і за всіма параметрами перевершує розглянуті моделі. Запропоновано, при побудові майбутніх моделей використовувати інші методи машинного вчення.

The article is devoted to the urgent problem of increasing the reliability of assessing the level of agricultural production. The work analyzed the nature of the existing dependence on the level of agricultural production of Ukraine. As factors affecting the volume of agricultural production, the following were considered: engaged in agricultural production, the cost of fixed assets, sown areas of crops. As a result of the initial analysis, emissions were identified. After eliminating emissions, a correlation matrix was built, which made it possible to identify a strong impact on the volume of agricultural production of the number of employees in it and the value of fixed assets. The third factor showed the strength of the connection at the border weak and medium. A system of linear and nonlinear both univariate and multifactorial regression models was constructed. The analysis of p-values for the internal parameters of the models allowed to exclude models with insignificant parameters from further consideration. The study of the values of the average approximation error of the constructed models revealed their exceeding the boundary value, which indicates the impossibility of using the obtained models to predict the volume of agricultural production. A decision was made for further attempts to build multifactorial regression models to use machine learning methods with a teacher: k-nearest neighbors and random forest. When comparing the obtained models, it was found that in both models the Fisher's calculated value exceeds the tabular values. Also, both models have quite high values of the determination coefficient. According to the results of the study of the values of the average approximation error, the model of k-nearest neighbors was discarded, so the value of the specified indicator for this

model exceeded the limit value. The random forest model has a valid approximation error value and is superior in all respects to the rest of the models considered. The obtained multivariate regression model meets all the initial requirements and can be recommended for assessing the level of agricultural production in Ukraine. It is proposed to use other methods of machine learning when building future models.

Ключові слова: *аграрне виробництво; регресія; модель k-найближчих сусідів; випадковий ліс; машинне навчання.*

Keywords: *agricultural production; regression; k-nearest neighbor model; random forest; machine learning.*

Постановка проблеми. Важливим показником, який характеризує ефективність діяльності сільськогосподарської галузі України є обсяг аграрного виробництва. Кризові явища в економіці країни негативно проявляють себе у стабільності роботи агропідприємств, підвищують ступінь ризику та невизначеність, що ще більше знижує інвестиційну привабливість підприємств галузі та їх кредитні можливості. Планування розвитку аграрного виробництва неможливо без дослідження кількісних факторів, які оказують вплив на рівень його обсягу.

Аналіз останніх досліджень і публікацій. Проблеми сільськогосподарської галузі та шляхи їх вирішення хвилювали багатьох дослідників, серед яких можна згадати роботи: Беженар І.М. [1], Дугієнко Н.О. [5], Духницький Б.В. [6], Загребельна І.Л. [7], Петлюк Л.О. [9], Симоненко Л.І. [10]. Безпосередньо питаннями аналізу факторів які впливають на обсяг аграрного виробництва та прогнозування рівня його обсягу присвячені праці Безпарточний М.Г. [2], Волошанюк Н.В. [3], Гадзало Я.М. [4], Лагодієнко Н.В. [8]. Проте дослідження кількісних факторів аграрного виробництва та проблема достовірності оцінювання його обсягу потребує свого розв'язання.

Постановка завдання. Метою дослідження є підвищення достовірності оцінювання обсягу аграрного виробництва України.

Виклад основного матеріалу. Початковими даними для проведення дослідження є динаміка обсягу аграрного виробництва України (Y) яка наведена у роботі Лагодієнко Н.В. [8, с. 45]. Факторами які досліджуються є: X1 - зайняті у сільськогосподарському виробництві, тис. осіб; X2 - вартість основних засобів, млн. грн.; X3 - посівні площі сільськогосподарських культур, тис. га. З метою перевірки початкових даних на викиди був побудований графік наведений на рис. 1. Графік показує, що у залежній змінній є один викид, у фактору X2 - два викиди та у фактору X3 - теж один викид.

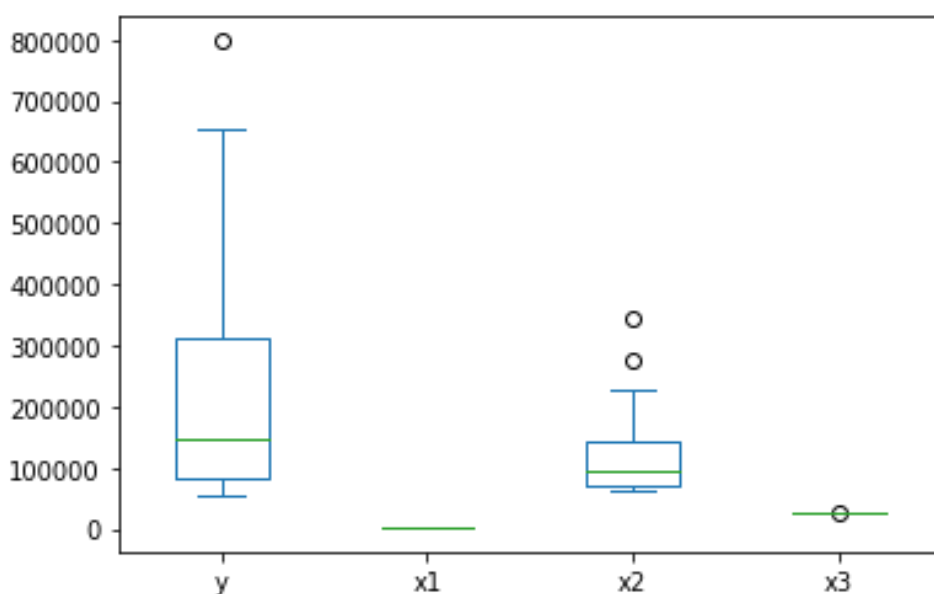


Рис. 1. Графік перевірки початкових даних на викиди

Після усунення викидів, з метою відбору факторів для наступних етапів дослідження була побудована кореляційна матриця яка наведена у табл. 1.

Таблиця 1. Кореляційна матриця

	Y	X1	X2	X3
Y	1	-0,771890586	0,945741618	0,499902697
X1	-0,771890586	1	-0,697804865	-0,318568874
X2	0,945741618	-0,697804865	1	0,592383424
X3	0,499902697	-0,318568874	0,592383424	1

Величини коефіцієнтів парних кореляцій (див. табл. 1) свідчать, що обсяг аграрного виробництва має дуже сильний зв'язок із вартістю основних засобів (величина коефіцієнта кореляції Пірсона $0,95 > 0,9$). Зайняті у сільськогосподарському виробництві мають сильний зворотній зв'язок з залежною змінною ($-0,77$). Фактор X_3 знаходиться за шкалою Чеддока на межі між слабким та середнім кореляційним зв'язком із залежною змінною. Тому, даний фактор був виключений із розгляду на подальших етапах дослідження. Необхідно вказати на наявність мультиколінеарності між факторами X_1 та X_2 (коефіцієнт парної кореляції приблизно дорівнює $-0,7$). З метою вирішення можливо чи ні побудувати двохфакторну модель була розрахована зворотна матриця для вказаних факторів (див. табл. 2).

Таблиця 2. Зворотна матриця для факторів X_1, X_2, R^{-1}

1,949058	1,360062
1,360062	1,949058

Інформація табл. 2 показує, що фактори X_1 та X_2 мають допустиму величину кореляційної залежності (величина розрахована на головній діагоналі зворотної матриці менше 4,0). За результатами аналізу кореляційних зав'язків було прийнято рішення побудувати лінійні та нелінійні регресійні моделі за формулами 1-3 та багатфакторні моделі за формулами 4-7.

$$\hat{y} = a_0 + a_1 \times x + \varepsilon, \quad (1)$$

$$\hat{y} = a_0 + a_1 \times x + a_2 \times x^2 + \varepsilon, \quad (2)$$

$$\hat{y} = a_0 \times x^{a_1} + \varepsilon, \quad (3)$$

$$\hat{y} = a_0 + a_1 \times x_1 + a_2 \times x_2 + \varepsilon, \quad (4)$$

$$\hat{y} = a_1 \times x_1 + a_2 \times x_2 + \varepsilon, \quad (5)$$

$$\hat{y} = a_0 \times x^{a_1} \times x^{a_2} + \varepsilon, \quad (6)$$

$$\hat{y} = x^{a_1} \times x^{a_2} + \varepsilon, \quad (7)$$

Розраховані параметри побудованих моделей наведені у табл. 3. З метою визначення значимості отриманих параметрів регресійних моделей був

використаний розрахунок p -значень для кожного з внутрішніх параметрів моделі (див. табл. 4).

Таблиця 3. Параметри моделей

Номер моделі	Параметри моделей		
	a_0	a_1	a_2
для фактору X1			
1	800017,27	-175,70	-
2	136142,08	200,99	-0,05
3	$10^{19,58}$	-4,063	-
для фактору X1			
1	-129619,35	3,063	-
2	83204,43	-1,398	2,12E-05
3	0,0003247	1,738	-
для факторів X1, X2			
4	97635,45	-49,67	2,57
5	-	-29,338	2,816
6	$10^{6,27}$	-1,95	1,18
7	-	-0,793	1,604

Таблиця 4. p -значення

Номер моделі	p -значення для параметру:		
	a_0	a_1	a_2
для фактору X1			
1	0,000110613	0,00074789	-
2	0,939591273	0,843357246	0,711470365
3	7,89164E-06	0,000156869	-
для фактору X2			
1	0,000670732	1,02466E-07	-
2	0,52691235	0,604641964	0,113692889
3	0,01661925	1,28021E-05	-
для факторів X1, X2			
4	0,440952411	0,081644981	1,62408E-05
5	-	0,000161517	4,00446E-09
6	0,142396029	0,025625178	0,002247081
7	-	0,002986207	1,32089E-07

Базуючись на величинах p -значень для кожного з внутрішніх параметрів були виключені із подальшого розгляду моделі де p -значення було більше за 0,05. Вважається, що якщо p -значення перевищує вказану величину, то параметри моделі ні є значимими. За результатами дослідження були відкинуті

дві однофакторні моделі та дві двохфакторні. Для порівняння моделей, які залишились на попередньому етапі, була сформована порівняльна табл. 5.

Таблиця 5. Порівняння моделей

Номер моделі	Коефіцієнт детермінації, R ²	Критерій Фішера, F		Середня помилка апроксимації, A
		розрахунковий	табличний	
для фактору X1				
1	0,595815077	19,163	4,667	28,057
3	0,67950495	27,562	4,667	26,155
для фактору X2				
1	0,894427209	110,138	4,667	27,16
3	0,780090822	46,115	4,667	23,423
для факторів X1, X2				
5	0,976862504	274,429	3,806	25,356
7	0,999551773	14495,077	3,806	21,191

Величини коефіцієнту детермінації (див. табл. 5) показує, що обидві моделі за фактором X1 та одна модель за фактором X2 є прийнятними ($0,5 < R^2 < 0,8$). Інші є достатньо хорошими, так як коефіцієнт кореляції більше за 0,8. Всі моделі є статистично значимими так як розрахунковий критерій Фішера більше табличного значення. Нажаль, незважаючи на вказані результати, значення середньої помилки апроксимації за всіма моделями перевищують граничне значення (10%). Тому, знайдені моделі для прогнозування обсягу аграрного виробництва використовувати неможливо.

За результатами попереднього етапу дослідження було прийнято рішення спробувати побудувати багатофакторну модель використовуючи методи машинного навчання. Серед методів машинного навчання було вирішено скористатися методами: k-ближніх сусідів та випадковим лісом.

Гіперпараметром моделі k-ближніх сусідів є кількість сусідніх елементів значення яких буде визначати прогнозну величину залежної змінної. Значення даного гіперпараметра у розраховані моделі дорівнює 3.

Основою для побудови моделі випадкового лісу є ансамбль слабких моделей, у нашому випадку використовуються бінарні дерева рішень. Гіперпараметрами побудованої регресійної моделі на основі випадкового лісу є:

глибина дерева, рівень точності приросту інформативності, кількість неоднакових предикатів у вузлу дерева. Критерієм якості під час поділу вузла був використаний квадрат помилки. Кількість дерев рішень у моделі 100. Приклад одного дерева рішень із моделі випадкового лісу наведений на рис. 2.

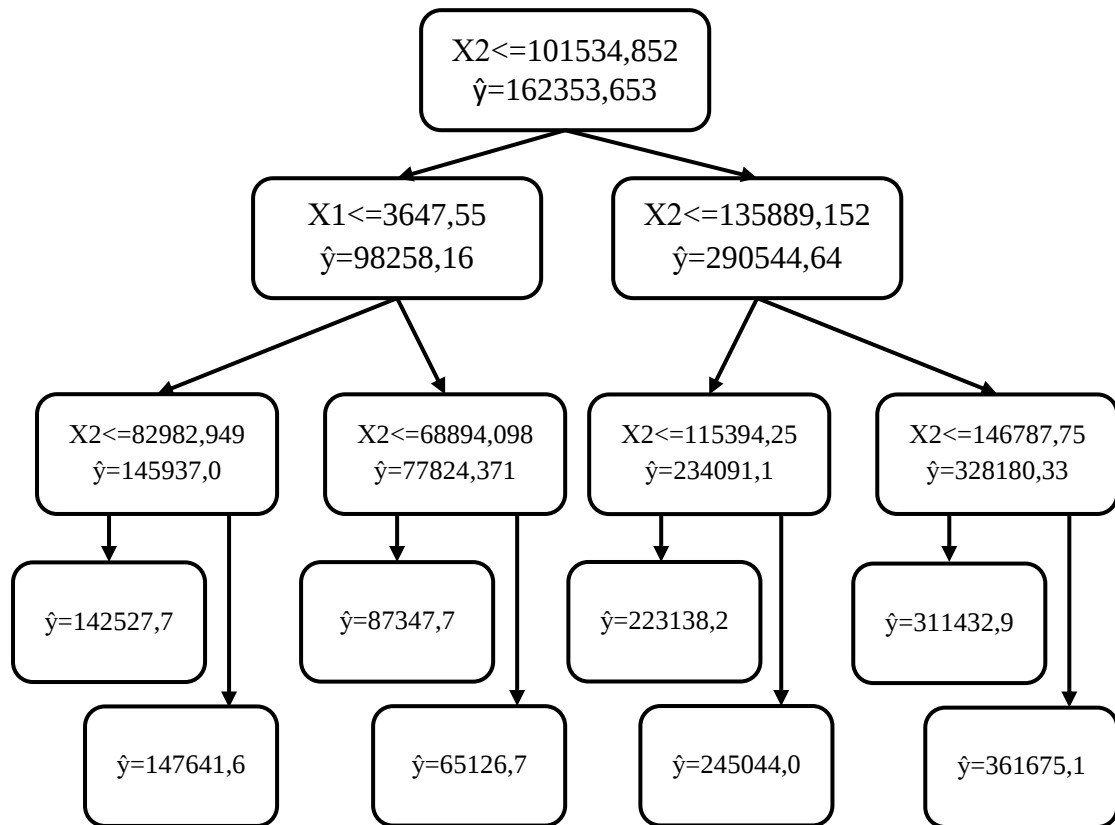


Рис. 2. Приклад одного з дерев рішень із моделі випадкового лісу

Для реалізації алгоритмів для побудови та навчання моделей машинного навчання була використана мова python та бібліотека Scikit-learn. Порівняння отриманих моделей наведена у табл. 6. Можна побачити, що обидві моделі мають достатньо добрі коефіцієнти детермінації. Для моделі випадкового лісу він перебільшую 0,95, що говорить про високу точність апроксимації. Розрахункові величини критерія Фішера для обох моделей також перевищують табличні значення.

Таблиця 6. Порівняння моделей машинного навчання для факторів X1, X2

Модель	Коефіцієнт детермінації, R ²	Критерій Фішера, F		Середня помилка апроксимації, A
		розрахунковий	табличний	
k-ближніх сусідів	0,929	78,855	3,806	16,30
випадковий ліс	0,979	278,428	3,806	8,20

Якщо, розглянути середню помилку апроксимації, то бачимо, що для моделі k-ближніх сусідів він складає більше 10%. Це говорить нам про те, що дану модель потрібно відкинути. Модель випадкового лісу задовольняє умові про точність апроксимації. Дану модель можна використовувати для оцінювання обсягу аграрного виробництва України. Модель можна представити у вигляді наступної формули:

$$\hat{y} = \frac{1}{n} \sum f_i(x_1, x_2) + \varepsilon, \quad (8)$$

Висновок. Розроблені моделі прогнозування рівня обсягу аграрного виробництва України. Були запропоновані лінійні та нелінійні однофакторні та багатфакторні регресійні моделі. Як альтернатива до вказаних моделей розроблені багатфакторні моделі за допомогою методів k-ближніх сусідів та випадкового лісу. В наслідок порівняння запропонованих моделей була відібрана найліпша - багатфакторна модель на основі випадкового лісу, де в якості факторів використовуються зайняті у сільськогосподарському виробництві та вартість основних засобів.

Перспективи подальших досліджень полягають у можливості підвищити якість оцінки рівня обсягу аграрного виробництва за рахунок застосування інших методів машинного навчання, наприклад, за рахунок використання нейронних мереж або XGBoost.

Література

1. Беженар І.М. Стан розвитку спеціалізації в аграрних підприємницьких структурах. *Глобальні та національні проблеми економіки*. 2016. № 10.

- С. 128-133. URL: <http://global-national.in.ua/archive/10-2016/28.pdf> (Дата звернення: 16.06.2022).
2. Безпарточний М.Г., Березіна Л.М., Баган Н.В. Витратний складник у стратегічному розвитку аграрних підприємств. *Проблеми системного підходу в економіці*. 2019. № 2(70). С. 57-64. URL: http://www.psae-jrnl.nau.in.ua/journal/2_70_1_2019_ukr/8.pdf (Дата звернення: 16.06.2022).
 3. Волошанюк Н.В., Сьомка А.К. Аграрний сектор України: тенденції, суб'єкти, перспективи реформування. *Науковий вісник Ужгородського національного університету*. 2017. № 13, ч. 1. С. 45-49. URL: http://www.visnyk-econom.uzhnu.uz.ua/archive/13_1_2017ua/12.pdf (Дата звернення: 16.06.2022).
 4. Гадзало Я.М., Лузан Ю.Я. Сучасні виклики модернізації аграрної політики Україн. *Вісник аграрної науки*. 2017. № 11. С. 9-22. URL: https://agrovisnyk.com/pdf/ua_2017_11_02.pdf (Дата звернення: 16.06.2022).
 5. Дугієнко Н.О., Лева В.Е. Сучасні тенденції розвитку аграрного сектору України. *Інфраструктура ринку*. 2018. № 25. С. 99-107. URL: http://www.market-infr.od.ua/journals/2018/25_2018_ukr/19.pdf (Дата звернення: 16.06.2022).
 6. Духницький Б.В. Світове виробництво сільськогосподарської продукції. *Економіка АПК*. 2019. № 7. С. 59-65. URL: http://eaprk.org.ua/sites/default/files/eaprk/2019/07/eaprk_2019_7_p_59_65.pdf (Дата звернення: 16.06.2022).
 7. Загребельна І.Л., Волкова Н.В., Олійник Є.О. Роль фермерських господарств в аграрному виробництві полтавської області. *Підприємництво та інновації*. 2019. № 10. С. 46-51. URL: <http://ei-journal.in.ua/index.php/journal/article/view/231/221> (Дата звернення: 16.06.2022).
 8. Лагодієнко Н.В. Моделювання кількісних параметрів впливу ключових факторів зростання аграрного виробництва. *Причорноморські економічні*

студії. 2020. № 51. С. 43-51. URL:
http://bses.in.ua/journals/2020/51_2020/9.pdf (Дата звернення: 16.06.2022).

9. Петлюк Л.О., Мєдвєдкова Н.С. Державна підтримка у забезпеченні розвитку аграрного сектору економіки України. *Економіка та держава*. 2021. № 2. С. 105-111. URL: http://www.economy.in.ua/pdf/2_2021/20.pdf (Дата звернення: 16.06.2022).
10. Симоненко Л.І., Захаріна О.В., Сайкевич М.І. Управління структурою виробництва в аграрному секторі. *Ефективна економіка*. 2018. № 2. URL: http://www.economy.nayka.com.ua/pdf/2_2018/62.pdf (Дата звернення: 16.06.2022).

References

1. Bezhenar, I.M. (2016), "State of specialization development in agricultural business structures", *Hlobalni ta natsionalni problemy ekonomiky*, vol. 10, pp. 128-133, available at: <http://global-national.in.ua/archive/10-2016/28.pdf> (Accessed 16 Jun 2022).
2. Bezpartochnyi, M., Berezina, L. and Bahan, N. (2019), "Cost component in strategic development of agricultural enterprises", *Problemy systemnoho pidkhodu v ekonomitsi*, vol. 2, pp. 57-64, available at: http://www.psae-jrnl.nau.in.ua/journal/2_70_1_2019_ukr/8.pdf (Accessed 16 Jun 2022).
3. Voloshanyuk, N.V. and Syomka, A.K. (2017), "Agricultural sector of Ukraine: Trends, business prospects of reform", *Naukovyi visnyk Uzhhorodskoho natsionalnoho universytetu*, vol. 13, no. 1, pp. 45-49, available at: http://www.visnyk-econom.uzhnu.uz.ua/archive/13_1_2017ua/12.pdf (Accessed 16 Jun 2022).
4. Hadzalo, Ya.M. and Luzan, Yu.Ya. (2017), "Modern challenges of modernization of agrarian policy of Ukraine", *Visnyk ahrarnoi nauky*, vol. 11, pp. 9-22, available at: https://agrovisnyk.com/pdf/ua_2017_11_02.pdf (Accessed 16 Jun 2022).
5. Dugienko, N.O. and Leva, V.E. (2018), "Modern trends in the development of the agricultural sector of Ukraine", *Infrastruktura rynku*, vol. 25, pp. 99-107,

available at: http://www.market-infr.od.ua/journals/2018/25_2018_ukr/19.pdf
(Accessed 16 Jun 2022).

6. Dukhnytskyi, B.V. (2019), "World agricultural production", *Ekonomika APK*, vol. 7, pp. 59-65, available at: http://eapk.org.ua/sites/default/files/eapk/2019/07/eapk_2019_7_p_59_65.pdf
(Accessed 16 Jun 2022).
7. Zahrebelna, I., Volkova, N. and Oliinyk, Ye. (2019), "The role of farming in agriculture in Poltava region" *Pidpryemnytstvo ta innovatsii*, vol. 10, pp. 46-51, available at: <http://ei-journal.in.ua/index.php/journal/article/view/231/221>
(Accessed 16 Jun 2022).
8. Lagodiienko, N. (2020), "Modeling of quantitative parameters of influence of key factors of agrarian production", *Prychornomorski ekonomichni studii*, vol. 51, pp. 43-51, available at: http://bses.in.ua/journals/2020/51_2020/9.pdf (Accessed 16 Jun 2022).
9. Petliuk, L. and Miedviedkova, N. (2021), "State support in ensuring the development of the agricultural sector of the Ukrainian economy", *Ekonomika ta derzhava*, vol. 2, pp. 105-111, available at: http://www.economy.in.ua/pdf/2_2021/20.pdf (Accessed 16 Jun 2022).
10. Symonenko, L., Zakharina, O. and Saikevych, M. (2018), "Managing production pattern in the agricultural sector", *Efektivna ekonomika*, vol. 2, available at: http://www.economy.nayka.com.ua/pdf/2_2018/62.pdf (Accessed 16 Jun 2022).

Стаття надійшла до редакції 24.06.2022 р.