

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ ВАДИМА ГЕТЬМАНА**

Інститут інформаційних технологій в економіці

Кафедра штучного інтелекту, моделювання та статистики

Освітньо-професійна програма «Економічна кібернетика»

Галузь знань 05 «Соціальні та поведінкові науки»

Спеціальність 051 «Економіка»

Форма навчання: _____ денна _____ очна

(денна), заочна, дистанційна

КВАЛІФІКАЦІЙНА БАКАЛАВРСЬКА РОБОТА

на тему «Математичне моделювання та аналіз сталого розвитку агровиробництва»

здобувача _____ Власова Дмитра Андрійовича _____

(ПІБ, підпис)

_____ (підпис)

Науковий керівник: к. е. н., доц. Кмитюк Тетяна Леонідівна

(науковий ступінь, учене звання, ПІБ)

_____ (підпис)

**Робота допущена до захисту перед екзаменаційною комісією з
атестації здобувачів вищої освіти (ЕК)**

Завідувач кафедри: к.ф.-м.н., проф. Великоіваненко Г.І.

(науковий ступінь, учене звання, ПІБ)

_____ (підпис)

Київ 2026

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ ВАДИМА ГЕТЬМАНА**

Інститут інформаційних технологій в економіці

Кафедра штучного інтелекту, моделювання та статистики

Освітньо-професійна програма «Економічна кібернетика»

Галузь знань 05 «Соціальні та поведінкові науки»

Спеціальність 051 «Економіка»

ПОГОДЖЕНО

Керівник проектної групи (гарант)
освітньо-професійної програми
Ольга ПРИТОМАНОВА

(підпис)

2026 р.

ЗАТВЕРДЖУЮ

Завідувач кафедри математичного
моделювання та статистики
Галина

ВЕЛИКОІВАНЕНКО

(підпис)

2026 р.

ІНДИВІДУАЛЬНЕ ЗАВДАННЯ

здобувача вищої освіти

Власов Дмитро Андрійович

(прізвище, ім'я, по батькові)

очної (денної)

форми здобуття освіти

очної (денної), заочної, дистанційної

на підготовку кваліфікаційної бакалаврської роботи

на тему

«Математичне моделювання та аналіз сталого розвитку агровиробництва»

Тему затверджено наказом ректора Університету від « 16 » березня 2026 р. № 477-ст

Кваліфікаційна бакалаврська робота виконується на матеріалах

сайтів Державної служби статистики <https://stat.gov.ua/> , Національного банку України
<https://bank.gov.ua/>

План кваліфікаційної бакалаврської роботи

Розділ 1	Теоретичні та методичні основи моделювання сталого розвитку агровиробництва
Розділ 2	Оцінювання сталого розвитку агровиробництва методами математичного моделювання та кластерного аналізу

Об'єкт дослідження:	агровиробництво України.
Предмет дослідження:	методи та моделі оцінювання та прогнозування сталого розвитку агровиробництва з урахуванням його структурних та часових характеристик.
Мета виконання кваліфікаційної бакалаврської роботи:	розроблення і практична апробація математичних моделей оцінювання сталого розвитку агровиробництва України на основі статистичного аналізу та сучасних методів моделювання.

Конкретні завдання, які здобувач повинен виконати для досягнення поставленої мети:

У розділі 1: 1) визначення сутності агровиробництва та теоретичні засади його сталого розвитку; 2) дослідження сучасних теоретико-методичних підходів оцінювання його сталого розвитку; 3) обґрунтування вибору методів економіко-математичного моделювання, машинного навчання та кластерного аналізу.

У розділі 2: 1) формування інформаційної бази дослідження на основі офіційних джерел; 2) проведення статистичного аналізу даних та дослідження їх структурних і часових особливостей; 3) побудова та порівняння моделей прогнозування обсягів агровиробництва для різних сегментів; 4) виконання кластеризації для виявлення структурної неоднорідності агровиробництва.

Завдання підготував
науковий керівник

(підпис)

Тетяна КМИТЮК

(ім'я, прізвище)

«__» _____ 2026 р.

Завдання одержав
здобувач

(підпис)

Дмитро ВЛАСОВ

(ім'я, прізвище)

«__» _____ 2026 р.

Реферат

Кваліфікаційна бакалаврська робота містить 64 сторінок, 18 таблиць, 20 рисунків, перелік використаних джерел з 33 найменувань.

«Математичне моделювання та аналіз сталого розвитку агровиробництва»

Об'єктом дослідження є агровиробництво України.

Предмет дослідження є економіко-математичні методи та моделі оцінювання сталого розвитку агровиробництва, а також його структурні та часові характеристики

Мета і завдання дослідження. Метою кваліфікаційної бакалаврської роботи є розроблення і практична апробація математичних моделей для оцінювання сталого розвитку агровиробництва України на основі статистичного аналізу та сучасних методів моделювання.

Для досягнення мети потрібно вирішити такі завдання:

1. Визначити сутність агровиробництва та теоретичні засади його сталого розвитку.
2. Дослідити сучасні теоретико-методичні підходи до оцінювання сталого розвитку в аграрному секторі.
3. Обґрунтувати вибір методів економіко-математичного моделювання, машинного навчання та кластерного аналізу.
4. Сформувати та підготувати статистичний датасет на основі офіційних джерел.
5. Провести статистичний аналіз даних та дослідити їх структурні і часові особливості.
6. Побудувати та порівняти моделі прогнозування обсягів агровиробництва для різних сегментів.
7. Виконати кластеризацію для виявлення структурної неоднорідності агровиробництва.

У процесі дослідження використано комплекс загальнонаукових та спеціальних методів, зокрема статистичний аналіз (описові статистики, кореляційний аналіз, ANOVA, STL-декомпозиція), економіко-математичне моделювання (моделі ETS та ARIMAX), методи машинного навчання (XGBoost, NNETAR) та методи кластерного аналізу (метод ліктя, k-means, ієрархічна кластеризація). Застосування зазначених методів дозволило виявити суттєву неоднорідність структури агровиробництва, наявність сезонності, трендових компонент і випадкових коливань у часових рядах, а також оцінити точність прогнозування для різних сегментів.

Теоретична, методична та практична значущість отриманих результатів.

Теоретична значущість роботи полягає в розвитку підходів до оцінювання сталого розвитку агровиробництва на основі поєднання статистичних, економетричних та машинно-навчальних методів.

Методична значущість полягає у формуванні комплексного інструментарію аналізу, який дозволяє досліджувати як структурні, так і часові особливості агровиробництва.

Практична значущість отриманих результатів полягає у можливості використання розроблених моделей для:

- підвищення точності прогнозів агровиробництва;
- оцінювання рівня сталого розвитку окремих сегментів;
- підтримки управлінських рішень у аграрному секторі.

Рік виконання кваліфікаційної бакалаврської роботи – 2026.

Рік захисту роботи – 2026.

Ключові слова: агровиробництво, сталий розвиток, математичне моделювання, прогнозування, кластеризація, машинне навчання, статистичний аналіз

ЗМІСТ

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ.....	3
ВСТУП.....	4
РОЗДІЛ 1	ТЕОРЕТИЧНІ
ТА МЕТОДИЧНІ ОСНОВИ МОДЕЛЮВАННЯ СТАЛОГО РОЗВИТКУ	ТА МЕТОДАМИ
АГРОВИРОБНИЦТВА	7
1.1. Сутність агровиробництва та особливості його розвитку в Україні.....	7
1.2. Методи прогнозування в дослідженні агровиробництва формалізація моделей, одиниці вимірювання.....	12
1.3. Методи кластеризації та критерії оцінювання.....	16
РОЗДІЛ 2	ОЦІНЮВАННЯ
СТАЛОГО РОЗВИТКУ АГРОВИРОБНИЦТВА	МЕТОДАМИ
МАТЕМАТИЧНОГО МОДЕЛЮВАННЯ ТА КЛАСТЕРНОГО АНАЛІЗУ ...	20
2.1 Статистична характеристика датасету.....	20
2.2 Побудова та оцінювання моделей прогнозування обсягів агровиробництва	27
2.3. Кластеризація видів сільськогосподарської продукції	52
ВИСНОВКИ	58
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	60

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ

ETS — Exponential Smoothing;

ARIMAX — Autoregressive Integrated Moving Average with Exogenous Variables;

XGBoost — Extreme Gradient Boosting;

NNETAR — Neural Network AutoRegression;

ME — Mean Error;

RMSE — Root Mean Squared Error;

MAE — Mean Absolute Error;

MAPE — Mean Absolute Percentage Error;

MASE — Mean Absolute Scaled Error;

MPE — Mean Percentage Error;

STL — Seasonal-Trend Decomposition using Loess;

ANOVA — Analysis of Variance;

AIC — Akaike Information Criterion;

AICc — Corrected Akaike Information Criterion;

BIC — Bayesian Information Criterion;

σ^2 — дисперсія залишків;

r — коефіцієнт кореляції;

Total_Volume — обсяг реалізованої продукції, т;

Avg_Price — середня ціна реалізації, грн/т;

Total_Area — площа посівів, га;

Total_Fert — обсяг внесених добрив, т;

Total_Invest — капітальні інвестиції, млн грн;

Total_Macro — обсяг продукції сільського господарства у постійних цінах 2021 року, млн грн.

k-means — метод k-середніх

ВСТУП

Актуальність теми: Агровиробництво є одною з основних галузей економіки України, що гарантує продовольчу безпеку та формує значну частину експортних надходжень. Водночас сучасний стан аграрного сектору характеризується підвищеною складністю умов функціонування, що пов'язано як із ресурсними обмеженнями, так і з впливом зовнішніх дестабілізуючих чинників, зокрема воєнних подій, нестабільності ринків та екологічних ризиків.

У практиці господарювання тривалий час домінував підхід, орієнтований на максимізацію виробництва та прибутковості, що часто супроводжувався інтенсивним використанням земельних ресурсів, збільшенням обсягів мінеральних добрив і навантаженням на екосистеми. У зв'язку з цим актуалізується необхідність переходу до концепції сталого розвитку, яка передбачає узгодження економічної результативності, екологічної безпеки та соціальної стійкості аграрного сектору. Складність структури галузі, наявність часових коливань, сезонності, а також взаємодія численних факторів зумовлюють необхідність застосування формалізованих математичних підходів. Використання економіко-математичних моделей, методів машинного навчання та кластерного аналізу дозволяє більш точно виявити закономірності розвитку агровиробництва, оцінити його структурну неоднорідність і сформулювати обґрунтовані прогнози.

Аналіз останніх досліджень і публікацій. Тему сталого розвитку в аграрній сфері детально вивчали вітчизняні економісти, зокрема О. М. Бородіна та В. М. Трегобчук. У їхніх працях добре описана сама теорія економічної та екологічної безпеки. Проте більшість класичних підходів спираються переважно на теоретичні моделі. Водночас у більшості досліджень переважає теоретичний підхід до оцінювання сталості, тоді як прикладні аспекти аналізу на основі масивів реальних даних залишаються менш розвиненими. Сучасні умови розвитку інформаційних технологій та наявність відкритих статистичних баз даних створюють передумови для використання методів економіко-математичного

моделювання, машинного навчання та кластеризації для отримання більш об'єктивної оцінки стану агровиробництва. Саме тому актуальним є дослідження, спрямоване на поєднання теоретичних підходів із практичним аналізом на основі емпіричних даних.

Мета і завдання дослідження. Метою кваліфікаційної бакалаврської роботи є розроблення і практична апробація математичних моделей для оцінювання сталого розвитку агровиробництва України на основі статистичного аналізу та сучасних методів моделювання.

Для досягнення мети потрібно вирішити такі завдання:

8. Визначити сутність агровиробництва та теоретичні засади його сталого розвитку.

9. Дослідити сучасні теоретико-методичні підходи до оцінювання сталого розвитку в аграрному секторі.

10. Обґрунтувати вибір методів економіко-математичного моделювання, машинного навчання та кластерного аналізу.

11. Сформувати та підготувати статистичний датасет на основі офіційних джерел.

12. Провести статистичний аналіз даних та дослідити їх структурні і часові особливості.

13. Побудувати та порівняти моделі прогнозування обсягів агровиробництва для різних сегментів.

14. Виконати кластеризацію для виявлення структурної неоднорідності агровиробництва.

Об'єкт дослідження — агровиробництво України.

Предмет дослідження — методи та моделі оцінювання та прогнозування сталого розвитку агровиробництва з урахуванням його структурних та часових характеристик.

Методи дослідження. У процесі дослідження використано сукупність загальнонаукових і спеціальних методів, зокрема:

- статистичний аналіз (описові статистики, кореляційний аналіз, ANOVA, STL-декомпозиція);
- економіко-математичне моделювання (моделі ETS, ARIMAX);
- методи машинного навчання (XGBoost, NNETAR);
- методи кластеризації (k-means, ієрархічна кластеризація);
- порівняльний аналіз на основі показників точності прогнозу (RMSE, MAE, MAPE, MASE).

Теоретична значущість роботи полягає в розвитку підходів до оцінювання сталого розвитку агровиробництва на основі поєднання статистичних, економетричних та машинно-навчальних методів.

Методична значущість полягає у формуванні комплексного інструментарію аналізу, який дозволяє досліджувати як структурні, так і часові особливості агровиробництва.

Практична значущість отриманих результатів полягає у можливості використання розроблених моделей для:

- підвищення точності прогнозів агровиробництва;
- оцінювання рівня сталого розвитку окремих сегментів;
- підтримки управлінських рішень у аграрному секторі.

Інформаційна база дослідження. Основою для проведення дослідження та розрахунків стали офіційні статистичні матеріали Державної служби статистики України (щодо посівних площ, реалізації сільськогосподарської продукції, використання добрив, капітальних інвестицій тощо), дані Національного банку України, а також наукові праці вітчизняних вчених.

РОЗДІЛ 1

ТЕОРЕТИЧНІ ТА МЕТОДИЧНІ ОСНОВИ МОДЕЛЮВАННЯ СТАЛОГО РОЗВИТКУ АГРОВИРОБНИЦТВА

1.1. Сутність агровиробництва та особливості його розвитку в Україні

Агровиробництво відіграє стратегічну роль у національній економіці України, забезпечуючи продовольчу безпеку та розвиток інших галузей. Зокрема, як зазначається у науковій літературі, «виняткову роль у структурі економіки України слід відвести стратегічно важливій галузі — сільському господарству. Адже позитивна динаміка розвитку аграрного сектору виступає гарантом продовольчої безпеки населення країни» [4]. У широкому розумінні агровиробництво являє собою систему процесів, що включає виробництво, первинну переробку та реалізацію продукції сільського господарства як рослинницького, так і тваринницького походження. Його функціонування пов'язане з використанням земельних, трудових, матеріально-технічних, фінансових та інвестиційних ресурсів, а результати безпосередньо залежать від природно-кліматичних умов, сезонності, технологічного рівня та кон'юнктури ринку [9].

У сучасній економічній літературі агровиробництво розглядається не лише як процес створення сільськогосподарської продукції, а і як складна соціально-економічна система, у якій поєднуються виробничі, ринкові, екологічні та інституційні компоненти [2]. Такий підхід особливо важливий для України, де аграрний сектор традиційно посідає вагоме місце в структурі економіки, а сільське господарство виконує не лише господарську, а й стратегічну функцію— підтримання продовольчої безпеки держави, забезпечення стабільності внутрішнього ринку та збереження експортного потенціалу [33]. Агровиробництво

водночас є сферою, де найбільш виразно проявляється взаємодія економічної ефективності та ресурсних обмежень, адже досягнення високих результатів нерідко супроводжується інтенсивним використанням ґрунтів, добрив, водних ресурсів і капіталу [28].

Специфіка агровиробництва полягає в тому, що воно значною мірою залежить від біологічних і природних процесів [18]. На відміну від багатьох інших видів економічної діяльності, сільське господарство має виражений сезонний характер, а виробничий цикл часто є тривалим та пов'язаним із природною ритмікою розвитку рослин і тварин [11]. Це зумовлює нерівномірність обсягів виробництва та реалізації продукції протягом року, формує циклічність цінкових змін і створює додаткові ризики для товаровиробників [8]. Для окремих сегментів агровиробництва характерні виражені сезонні коливання, тоді як для інших більш притаманними є трендова динаміка або значна частка випадкових змін, що підтверджується результатами аналізу часової структури відповідних показників.

В українських умовах агровиробництво має чітко виражену галузеву структуру, до якої входять рослинництво і тваринництво. Рослинницький напрям охоплює виробництво зернових і зернобобових культур, кукурудзи, картоплі, овочевих культур, сої та інших видів продукції, тоді як тваринницький напрям включає вирощування великої рогатої худоби, свинарство, птахівництво та інші підгалузі. Кожен із цих сегментів відрізняється за виробничим циклом, рівнем капіталомісткості, ресурсною інтенсивністю, залежністю від цінової кон'юнктури та ринковими каналами збуту.

Для України агровиробництво має особливе значення завдяки поєднанню природно-ресурсних переваг і високого експортного потенціалу. Значні площі сільськогосподарських угідь, висока частка родючих чорноземів, сприятливі агрокліматичні умови та історично сформована спеціалізація на рослинницькій продукції створили передумови для формування потужного аграрного сектору [16]. У структурі українського агровиробництва значне місце займають зернові культури, передусім пшениця, кукурудза, ячмінь, а також олійні культури, зокрема соя [3]. Поряд із цим важливу роль відіграють картопля, овочева продукція та

продукція тваринництва, що забезпечують внутрішній споживчий ринок і формують продовольчу стабільність країни [7].

Агровиробництво України характеризується відсутністю однорідності як за масштабами діяльності, так і за організаційною структурою. У ньому співіснують великі агрохолдинги, фермерські господарства, сільськогосподарські підприємства та дрібні виробники. Розглянуті суб'єкти агровиробництва суттєво різняться за технічним рівнем, доступом до інвестицій, впровадженням інноваційних технологій, продуктивністю праці та здатністю протистояти ринковим ризикам [12]. Великотоварне виробництво, особливо в зерновому сегменті, вирізняється високою концентрацією ресурсів, значними посівними площами, інтенсивним використанням добрив та орієнтацією на зовнішні ринки [1]. Виробництво продукції для внутрішнього ринку, а також частина тваринницьких сегментів характеризуються меншими масштабами, діяльності, відмінною структурою витрат і нижчим рівнем капіталоемності [29]. У результаті агровиробництво в Україні доцільно розглядати як багатокомпонентну систему з різними моделями функціонування.

Сутність агровиробництва найбільш повно розкривається через систему показників, які характеризують його результативність, ресурсне забезпечення та економічні умови функціонування. У межах цієї роботи до ключових параметрів віднесено обсяги реалізації продукції, середні ціни реалізації, площі посівів, обсяги внесених добрив, капітальні інвестиції та обсяги продукції сільського господарства у постійних цінах 2021 року [5; 6]. Такий підхід дозволяє інтерпретувати агровиробництво не лише як фізичний випуск продукції, а як результат взаємодії виробничих ресурсів, інвестиційної активності та ринкового середовища. При цьому вже на етапі статистичного аналізу виявлено, що більшість відповідних показників характеризуються значною варіативністю, асиметрією розподілів, наявністю викидів та різною повнотою спостережень, що свідчить про суттєву неоднорідність аграрного середовища України.

Концепція сталого розвитку відіграє важливу роль у формуванні сучасного агровиробництва та визначає напрям його розвитку. У наукових дослідженнях

наголошується, що «соціально-економічний розвиток має ґрунтуватися на принципах врахування можливостей природних комплексів витримувати антропогенні навантаження і забезпечувати нормальне функціонування біосфери і локальних екосистем» [13]. Реалізація даного підходу спрямована на забезпечення збалансованого врахування економічних, екологічних і соціальних складових розвитку аграрного сектору. Тривалий час результативність аграрного сектору оцінювалася переважно через показники обсягів виробництва, урожайності, прибутковості та експорту. Однак такий підхід не враховує довгострокових наслідків надмірного використання природних ресурсів та впливу зовнішніх дестабілізуючих чинників [20]. Особливістю українського агровиробництва є те, що воно функціонує в умовах тривалої воєнної агресії Російської Федерації, яка розпочалася у 2014 році, а з 2022 року переросла у повномасштабну війну [10]. Це суттєво змінило умови ведення сільськогосподарської діяльності, особливо в регіонах, наближених до зони бойових дій. Воєнні дії призводять до скорочення оброблюваних і засіяних площ, оскільки частина сільськогосподарських угідь забруднена вибухонебезпечними предметами, а частина земель тимчасово вибуває з господарського використання через безпекові обмеження та наслідки бойових дій. За даними ФАО, загальна площа оброблюваних земель в Україні у 2024 році зменшилася на 7 % порівняно з 2023 роком, при цьому особливо гостро проблема проявляється у прифронтових регіонах [19].

Крім зменшення посівних площ, війна зумовлює суттєве погіршення стану ресурсної бази агровиробництва. Значні території сільськогосподарських угідь зазнають фізичного пошкодження, забруднення залишками боєприпасів і токсичними речовинами, що погіршує якість ґрунтів та знижує їхню продуктивність у довгостроковій перспективі [30]. Одночасно руйнується виробнича, складська, енергетична та транспортна інфраструктура, ускладнюється доступ до елеваторів, техніки, пального, електроенергії та каналів збуту, що безпосередньо знижує ефективність аграрного виробництва і підвищує його витрати [31]. За оцінками міжнародних інституцій, війна завдала українському агросектору десятків мільярдів доларів збитків і втрат, а окремим напрямом

післявоєнного відновлення стане гуманітарне розмінування сільськогосподарських земель [27].

Ще одним важливим наслідком війни для агровиробництва є дефіцит трудових ресурсів. У воєнних умовах аграрні підприємства стикаються з нестачею працівників через мобілізаційні процеси, міграцію населення, скорочення трудового потенціалу сільських територій та загальне зростання кадрового навантаження на бізнес. ФАО прямо відносить нестачу робочої сили до ключових проблем агропідприємств [19], а профільний парламентський комітет у 2025 році наголошував, що дефіцит трудових ресурсів у сільській місцевості на тлі війни вже створює ризики для продовольчої безпеки. У зв'язку з цим сучасне розуміння агровиробництва має виходити за межі суто виробничої логіки та включати критерії екологічної безпеки, відтворення ресурсної бази, стійкості до кризових шоків і довгострокової економічної рівноваги. Сталий розвиток агровиробництва передбачає гармонізацію трьох взаємопов'язаних складових—економічної, екологічної та соціальної [15]. Економічна складова пов'язана з підтриманням продуктивності, конкурентоспроможності та інвестиційної привабливості аграрного сектору; екологічна—з раціональним використанням земельних і природних ресурсів, недопущенням деградації ґрунтів і мінімізацією наслідків техногенного та воєнного забруднення; соціальна—із збереженням зайнятості, підтримкою сільських територій і забезпеченням продовольчої безпеки країни [21]. Саме тому в сучасних дослідженнях дедалі частіше трактується як система, ефективність якої має оцінюватися не лише за кінцевими обсягами продукції, а й за здатністю забезпечувати збалансований розвиток у довгостроковій перспективі навіть в умовах воєнних викликів [23].

Агровиробництво слід визначати як складну багаторівневу систему виробництва та реалізації сільськогосподарської продукції, що функціонує під впливом природних, економічних, технологічних, інституційних і воєнно-безпекових чинників та має вирішальне значення для продовольчої безпеки, експорту й економічної стабільності України. Його сутність полягає не лише у створенні матеріального продукту, а й у формуванні довгострокової економічної

основи розвитку країни. Саме тому дослідження агровиробництва в сучасних умовах повинно базуватися на поєднанні теоретичного осмислення його змісту з кількісними методами оцінювання, що дозволяють виявляти закономірності розвитку, неоднорідність окремих сегментів і потенційні ризики. Такий підхід формує теоретичне підґрунтя для подальшого розгляду методів прогнозування, кластеризації та моделювання сталого розвитку агровиробництва, що здійснюється у наступному підрозділі.

1.2. Методи прогнозування в дослідженні агровиробництва формалізація моделей, одиниці вимірювання

У сучасних дослідженнях агровиробництва дедалі більшого значення набувають методи математичного моделювання, що дають змогу не лише описувати наявні тенденції, а й формувати прогнози розвитку окремих сегментів аграрного сектору та виявляти однорідні групи об'єктів за сукупністю характеристик. У межах даної роботи для аналізу часової динаміки та прогнозування обсягів реалізації сільськогосподарської продукції використано моделі, що належать до різних класів: класичні статистичні моделі експоненційного згладжування ETS, економетричні моделі часових рядів із зовнішніми факторами ARIMAX, нейромережеві моделі NNETAR та алгоритми машинного навчання XGBoost. Для задач сегментації та групування об'єктів застосовано метод ліктя для визначення кількості кластерів, алгоритм k-means і метод ієрархічної кластеризації. Такий підхід дає змогу поєднати переваги класичного статистичного аналізу, економетрії, машинного навчання та багатовимірної класифікації [26].

Інформаційною основою дослідження є набір змінних, що характеризують виробничі, ресурсні, цінові, інвестиційні та макроекономічні аспекти функціонування аграрного сектору. До основних показників належать:

Total_Volume—обсяг реалізованої продукції, т; Avg_Price—середня ціна реалізації, грн/т; Total_Area—площа посівів, га; Total_Fert—обсяг внесених добрив, т; Total_Invest—капітальні інвестиції, млн грн; Total_Macro—обсяг продукції сільського господарства у постійних цінах 2021 року, млн грн. Обрані показники дають змогу комплексно дослідити обсяги агровиробництва, враховуючи ресурсні, ринкові та макроекономічні аспекти функціонування галузі.

Оскільки досліджувані показники мають різні масштаби вимірювання, під час побудови моделей та особливо під час кластеризації важливо враховувати необхідність попередньої стандартизації або нормалізації ознак. Це пояснюється тим, що показники, виражені у великих абсолютних значеннях (наприклад, гектари чи мільйони гривень), можуть домінувати в розрахунках відстаней і спотворювати результати кластеризації. Саме тому в багатовимірному аналізі доцільно перейти від вихідних значень до стандартизованих.

Моделі ETS (Error, Trend, Seasonality) належать до класу традиційних статистичних методів прогнозування часових рядів і є особливо корисними для даних, у яких присутні трендова та/або сезонна складові. Перевагами ETS-моделей є відносна простота побудови, інтерпретованість і здатність адаптуватися до змін динаміки ряду через механізм експоненційного згладжування. У дослідженні ETS використовується як базовий статистичний інструмент, що дозволяє відтворити часову структуру ряду без явного включення зовнішніх факторів.

У загальному вигляді для моделі з рівнем, трендом і сезонністю формалізація може бути подана так:

$$\begin{aligned} l_t &= \alpha(y_t - s_{t-m}) + (1 - \alpha)(l_{t-1} + b_{t-1}), \\ b_t &= \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1}, \\ s_t &= \gamma(y_t - l_t) + (1 - \gamma)s_{t-m}, \\ \hat{y}_{t+h} &= l_t + hb_t + s_{t-m+hm}, \end{aligned} \quad (1.1)$$

де y_t — фактичне значення ряду в момент часу t ; l_t — поточний рівень ряду; b_t — трендова складова; s_t — сезонна складова; m — сезонний період; α, β, γ — коефіцієнти згладжування.

Якщо тренд або сезонність відсутні, відповідні рівняння спрощуються. У випадку місячних даних сезонний цикл, як правило, формується протягом 12 періодів, що відповідає одному року [24].

ARIMAX є розширенням класичних моделей ARIMA, у яких поряд із внутрішньою часовою залежністю враховується вплив зовнішніх пояснювальних змінних. Для агровиробництва це особливо важливо, оскільки обсяги реалізації продукції залежать не лише від попередньої динаміки самого показника, а й від цін, площ посівів, обсягів добрив, інвестиційної активності та загальної макроекономічної ситуації. У цьому дослідженні ARIMAX використовується саме як економетричний інструмент поєднання часової структури ряду та факторного аналізу.

У загальному вигляді ARIMAX-модель можна записати так:

$$y_t = c + \sum_{i=1}^p \phi_i y_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \sum_{k=1}^m \beta_k x_{k,t} + \varepsilon_t, \quad (1.2)$$

де y_t — значення цільової змінної у момент часу t ; c — константа; ϕ_i — коефіцієнти авторегресії; θ_j — коефіцієнти ковзного середнього; $x_{k,t}$ — значення k -го зовнішнього фактора; β_k — параметри впливу екзогенних змінних; ε_t — випадкова похибка.

Для сезонних модифікацій до моделі додаються сезонні авторегресивні та сезонні ковзні компоненти у формі ARIMAX(p,d,q)(P,D,Q)[s], де s — довжина сезону [22].

NNETAR (Neural Network AutoRegression) належить до класу нейромережевих моделей для часових рядів. Її ключова перевага полягає у здатності виявляти складні нелінійні залежності, які не завжди можуть бути описані класичними статистичними методами. У контексті агровиробництва це є важливим, оскільки динаміка реалізації продукції часто формується під впливом нерівномірних сезонних змін, ринкових шоків і нелінійної взаємодії факторів. У роботі NNETAR розглядається як інструмент для моделювання складної структури рядів, особливо там, де лінійні моделі виявляються недостатньо гнучкими [25].

У формалізованому вигляді NNETAR може бути подана як нелінійна функція від лагових значень ряду:

$$y_t = f(y_{t-1}, y_{t-2}, \dots, y_{t-p}, y_{t-s}, \dots, y_{t-ks}, x_t) + \varepsilon_t, \quad (1.3)$$

де $f(\cdot)$ —нелінійна функція, яка апроксимується штучною нейронною мережею; p —кількість несезонних лагів; s —сезонний період; k —кількість сезонних лагів; x_t —можливі зовнішні пояснювальні змінні; ε_t —випадкова похибка.

Фактично модель навчається на історичних спостереженнях і намагається відтворити приховану залежність між попередніми та майбутніми значеннями ряду.

XGBoost (Extreme Gradient Boosting) належить до класу ансамблевих методів машинного навчання і базується на послідовному побудуванні дерев рішень, кожне з яких зменшує помилки попередніх. Перевагою цього підходу є висока гнучкість, здатність працювати з великою кількістю ознак, у тому числі лагових значень ряду, а також ефективність у задачах, де зв'язки між змінними мають складний і неоднорідний характер. У дослідженні XGBoost використовується як ключовий інструмент прогнозування, що поєднує часові лаги, факторні ознаки та механізм бустингу для підвищення точності прогнозів [17].

У загальному вигляді прогноз XGBoost подається як сума базових алгоритмів (дерев):

$$\hat{y}_i = \sum_{m=1}^M f_m(x_i), \quad f_m \in \mathcal{F}, \quad (1.4)$$

$\hat{y}_i$ —прогнозне значення для i -го спостереження; f_m —окреме дерево рішень; M —кількість дерев у моделі; $\mathcal{F}$ —простір усіх допустимих дерев; x_i —вектор ознак, який може включати як факторні змінні, так і лагові значення цільового показника.

Поступове додавання нових дерев дозволяє моделі зменшувати функцію втрат і враховувати складні нелінійні ефекти.

1.3. Методи кластеризації та критерії оцінювання

Окремим напрямом дослідження є кластерний аналіз, який дає змогу групувати об'єкти за схожістю їх характеристик. У даній роботі кластеризація використовується для виявлення однорідних груп об'єктів за показниками, що характеризують агровиробництво, ресурсне забезпечення, ціни, добрива, площі посівів та інші параметри. Для цього було застосовано поєднання трьох підходів: метод ліктя для вибору кількості кластерів, алгоритм k-means для фактичного розбиття на кластери та ієрархічну кластеризацію для побудови структури подібності між об'єктами.

Метод ліктя застосовується для визначення оптимальної кількості кластерів k . Його суть полягає у послідовному обчисленні сумарної внутрішньокластерної дисперсії для різних значень k та виявленні точки, після якої додавання нових кластерів вже не забезпечує істотного покращення якості групування. У такій точці графік набуває форми «ліктя», що і дає назву методу. У дослідженні цей підхід використовується як попередній етап перед запуском k-means.

Внутрішньокластерна сума квадратів може бути записана так:

$$WCSS(k) = \sum_{r=1}^k \sum_{x_i \in C_r} \|x_i - \mu_r\|^2, \quad (1.5)$$

де C_r — r -й кластер; μ_r — центр кластера; $\|x_i - \mu_r\|^2$ — квадрат евклідової відстані між об'єктом i і центром кластера.

Оптимальне значення k вибирається в точці, де темп зменшення WCSS різко сповільнюється.

Алгоритм k-means належить до методів несупервізованого навчання та використовується для групування об'єктів у задану кількість кластерів. Його мета — мінімізувати суму квадратів відстаней між об'єктами та центрами відповідних кластерів. Остаточне групування об'єктів у кластери здійснюється за допомогою

алгоритму k-means на основі попередньо визначеного оптимального значення k методом ліктя [32].

Функція мети алгоритму має вигляд:

$$\min \sum_{r=1}^k \sum_{x_i \in C_r} \|x_i - \mu_r\|^2, \quad (1.6)$$

де C_r — множина об'єктів r -го кластера, а μ_r — його центр.

Алгоритм ітеративно виконує два основні кроки: призначає кожен об'єкт до найближчого центру та оновлює координати центрів як середні значення всіх об'єктів кластера. Процес повторюється до досягнення збіжності.

Ієрархічна кластеризація є альтернативним підходом до сегментації, який дозволяє побудувати багаторівневу структуру подібності між об'єктами. Її результат зазвичай подається у формі дендрограми — графічного зображення, схожого на гілки дерева, де поступово поєднуються найбільш подібні об'єкти або кластери. У дослідженні використано агломеративний підхід, за якого спочатку кожен об'єкт вважається окремим кластером, а далі найближчі кластери послідовно об'єднуються до формування єдиної ієрархічної структури. У програмній реалізації застосовано евклідову відстань та правило об'єднання Ward.D2, що орієнтоване на мінімізацію приросту внутрішньокластерної дисперсії [14].

Загальна логіка ієрархічної кластеризації базується на матриці відстаней:

$$d_{ij} = \sqrt{\sum_{m=1}^p (x_{im} - x_{jm})^2}, \quad (1.7)$$

де d_{ij} — евклідова відстань між об'єктами i та j ; x_{im} , x_{jm} — значення m -ї ознаки для відповідних об'єктів.

Критерії оцінювання прогнозних моделей

Для забезпечення коректного порівняння моделей прогнозування в дослідженні використано єдину систему оцінювання якості прогнозу. Відповідно до організації експерименту часові ряди було поділено на навчальну і тестову вибірки у співвідношенні 80/20, де навчальна частина застосовувалася для

побудови моделі, а тестова — для перевірки її здатності узагальнювати закономірності поза межами періоду навчання. Для оцінювання використовувалися метрики ME, RMSE, MAE, MAPE та MASE, що дозволяє зіставляти моделі різних класів за єдиною системою критеріїв.

Середня помилка прогнозу визначається як:

$$ME = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i), \quad (1.8)$$

де y_i —фактичне значення, $\hat{y}_i$ —прогнозне значення, n —кількість спостережень.

Корінь із середньоквадратичної помилки визначається формулою:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (1.9)$$

де y_i —фактичне значення, $\hat{y}_i$ —прогнозне значення, n —кількість спостережень.

Середня абсолютна помилка визначається так:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (1.10)$$

де y_i —фактичне значення, $\hat{y}_i$ —прогнозне значення, n —кількість спостережень.

Середня абсолютна відносна помилка визначається так:

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (1.11)$$

де y_i —фактичне значення, $\hat{y}_i$ —прогнозне значення, n —кількість спостережень.

Середня абсолютна масштабована помилка визначається так:

$$MASE = \frac{\frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|}{\frac{1}{n-1} \sum_{t=2}^n |y_t - y_{t-1}|} \quad (1.12)$$

де y_i —фактичне значення, $\hat{y}_i$ —прогнозне значення, n —кількість спостережень

В межах даного дослідження застосовано комплексний інструментарій аналізу, який поєднує методи прогнозування та кластеризації. Прогнозний блок представлений чотирма підходами різної природи — ETS, ARIMAX, NNETAR та XGBoost, що дозволяє зіставити класичні статистичні, економетричні, неймережеві та машинно-навчальні підходи до прогнозування часових рядів. Кластерний блок охоплює метод ліктя, алгоритм k-means та ієрархічну кластеризацію, що дає змогу не лише визначати кількість груп, а й аналізувати внутрішню структуру подібності між об'єктами. Оскільки змінні мають різні одиниці вимірювання, виникає потреба у їх стандартизації. Використання системи метрик ME, RMSE, MAE, MAPE та MASE дає змогу здійснювати комплексне оцінювання точності прогнозування. Саме така методична база формує підґрунтя для подальшого практичного аналізу в розділі 2, де наведено результати застосування зазначених підходів до сформованого датасету агровиробництва.

РОЗДІЛ 2

ОЦІНЮВАННЯ СТАЛОГО РОЗВИТКУ АГРОВИРОБНИЦТВА МЕТОДАМИ МАТЕМАТИЧНОГО МОДЕЛЮВАННЯ ТА КЛАСТЕРНОГО АНАЛІЗУ

2.1 Статистична характеристика датасету

У межах дослідження було сформовано датасет на основі системи показників, обґрунтованих у підрозділі 1.2, що відображають основні характеристики агровиробництва. У цьому підрозділі основну увагу зосереджено на статистичній характеристиці масиву даних, зокрема на описових статистиках, формі розподілів, кореляційних зв'язках і часових особливостях досліджуваних показників.

Таблиця 2.1 – Статистична характеристика змінних сформованого датасету

Variable	n	mean	sd	median	trimmed	mad
Total_Volume	1 249	757 591	1526 197	81 924	375 098	98 954
Avg_Price	1 430	17 340.8	17406.2	8 138.1	14 445.1	6 913.9
Total_Area	1 152	9 796 173	12 674 744	2 508 550	7 862 649	3 705 758
Total_Fert	7 20	10 167.8	9 424.5	8 529.3	9 047	10 030.3
Total_Invest	1 872	12 616.1	12 876.9	6 819.7	10 040	7 739.6
Total_Macro	1 872	40 691.7	20 777.8	48 968.5	41 540.6	8 985

Джерело побудовано автором

Дані (табл. 2.1) свідчать, що кількість спостережень для окремих змінних є неоднаковою, сформований датасет характеризується різним ступенем повноти інформації за окремими показниками. Зокрема, для змінної Total_Fert кількість спостережень становить 720, для Total_Area — 1152, для Total_Volume — 1249, для

Avg_Price — 1430, тоді як для Total_Invest та Total_Macro зафіксовано по 1872 спостереження. Це дає підстави стверджувати, що частина змінних містить пропуски або була сформована на основі менш повного масиву первинних даних, що необхідно враховувати на наступних етапах статистичної обробки. Водночас найбільш повно представленими у вибірці є показники капітальних інвестицій та обсягу продукції сільського господарства у постійних цінах, що створює кращі передумови для їх подальшого використання у порівняльному та модельному аналізі.

Результати аналізу описових статистик, зокрема середніх значень, медіан і стандартних відхилень, підтверджують наявність значних відмінностей між змінними як за масштабом, так і за рівнем варіативності. Для показника Total_Volume середнє значення становить 757591.3 т, тоді як медіана дорівнює лише 81924.2 т, а стандартне відхилення — 1526197.7, що свідчить про значну неоднорідність обсягів реалізації та наявність окремих великих значень у вибірці. Подібна ситуація спостерігається і для Total_Area, де середнє значення становить 9796173.9 га, медіана — 2508550 га, а стандартне відхилення — 12674744.3, що вказує на суттєву диференціацію об'єктів дослідження за масштабами посівних площ. Для змінної Avg_Price середнє значення становить 17340.8 грн/т, медіана — 8138.1 грн/т, а стандартне відхилення — 17406.2, що також відображає значну варіацію цін реалізації в межах вибірки. Показник Total_Fert характеризується середнім рівнем 10167.8 т, медіаною 8529.4 т і стандартним відхиленням 9424.5, тобто варіація цього показника є помітною, але менш різкою, ніж для обсягу реалізації чи площ посівів. Для Total_Invest середнє значення дорівнює 12616.1 млн грн, медіана — 6819.7 млн грн, а стандартне відхилення — 12876.9, що свідчить про істотну нерівномірність інвестиційної активності. Змінна Total_Macro має середнє значення 40691.7 млн грн, медіану 48968.5 млн грн і стандартне відхилення 20777.9, що дає підстави говорити про значну, але більш збалансовану варіативність порівняно з деякими іншими показниками. Загалом суттєві розбіжності між середніми значеннями та медіанами для більшості змінних свідчать про асиметричність розподілів, а високі значення стандартних відхилень

підтверджують неоднорідність досліджуваної сукупності, що є важливою характеристикою датасету перед проведенням подальшого багатовимірного аналізу.

Таблиця 2.2 – Показники форми розподілу та варіації досліджуваних змінних

Variable	min	max	range	skew	kurtosis	se
Total_Volume	15.7	19 931 190	19 931 175	3.7	24.9	43 184.6
Avg_Price	982.8	73 385.9	72 403.1	1.2	0.12	460.3
Total_Area	1 800	37 904 400	37 902 600	1	-0.52	373 433.2
Total_Fert	0.3	35 199.9	35 199.6	0.8	-0.2	351.2
Total_Invest	1 349.4	45 285.2	43 935.7	1.5	1.1	297.6
Total_Macro	9 709.7	67 080.6	57 370.9	-0.6	-1.3	480.2

Джерело: побудовано автором

Аналіз показників асиметрії та ексцесу, наведених у (табл. 2.2), свідчить про те, що розподіли більшості досліджуваних змінних не є симетричними, а окремі з них характеризуються наявністю виражених викидів. Для показника Total_Volume коефіцієнт асиметрії становить 3.7, а ексцес — 24.9, що вказує на різко виражену правосторонню асиметрію та надмірну концентрацію окремих екстремально великих значень у вибірці. Така форма розподілу означає, що в масиві даних присутня відносно невелика кількість спостережень з дуже високими обсягами реалізованої продукції, які істотно зміщують середнє значення у бік збільшення. Для змінних Avg_Price, Total_Area, Total_Fert та Total_Invest також зафіксовано додатні значення асиметрії — відповідно 1.2, 1.02, 0.8 та 1.5, що свідчить про правосторонню асиметрію, хоча менш виражену, ніж у показника обсягу реалізації. Водночас для Total_Macro спостерігається від’ємна асиметрія на рівні -0.6 та від’ємний ексцес -1.3, що характеризує розподіл як більш згладжений і зміщений у лівий бік порівняно з іншими показниками. Форма розподілів досліджуваних ознак є неоднорідною, а наявність значних значень асиметрії та ексцесу

підтверджує необхідність врахування викидів і відхилень від нормального розподілу в подальшому аналізі.

Таблиця 2.3 – Кореляційна матриця змінних датасету

Variable	Total_Volume	Avg_Price	Total_Area	Total_Fert	Total_Invest	Total_Macro
Total_Volume	1	-0.23	0.29	0.33	-0.15	0.2
Avg_Price	-0.23	1	-0.08	0.13	0.75	-0.1
Total_Area	0.29	-0.08	1	0.48	-0.14	0.07
Total_Fert	0.33	0.13	0.48	1	0.02	0.1
Total_Invest	-0.15	0.75	-0.14	0.02	1	-0.07
Total_Macro	0.2	-0.1	0.07	0.1	-0.07	1

Джерело: побудовано автором

Аналіз кореляційної матриці, наведеної в (табл. 2.3), дає змогу оцінити силу та напрям взаємозв'язків між основними показниками датасету, що є важливим етапом перед застосуванням математичних методів аналізу. Найтісніший позитивний зв'язок спостерігається між змінними Avg_Price та Total_Invest ($r = 0.75$), що може свідчити про те, що зростання середньої ціни реалізації супроводжується вищим рівнем капітальних інвестицій, або навпаки — більш інвестиційно забезпечені об'єкти мають змогу формувати вищу вартість реалізованої продукції. Помірний позитивний зв'язок також зафіксовано між Total_Area та Total_Fert ($r = 0.48$), що є економічно логічним, оскільки збільшення площ посівів зазвичай супроводжується зростанням обсягів внесення добрив. Крім того, показник Total_Volume має помірний додатний зв'язок із Total_Fert ($r = 0.33$) та Total_Area ($r = 0.29$), що вказує на наявність зв'язку між масштабами виробництва, інтенсивністю ресурсного забезпечення та обсягами реалізації продукції. Водночас між Total_Volume та Avg_Price спостерігається слабкий від'ємний зв'язок ($r = -0.23$), а між Total_Volume та Total_Invest — також слабкий від'ємний зв'язок ($r = -0.15$), що характеризує слабкий взаємозв'язок між обсягами реалізації, рівнем цін і інвестиційною активністю у вибірці. Інші коефіцієнти кореляції є переважно слабкими, зокрема зв'язок Total_Macro з рештою змінних коливається в межах від -0.10 до 0.20, що вказує на відносну автономність цього

показника в загальній структурі датасету. Загалом отримані результати дають підстави стверджувати, що між окремими змінними існують змістовні економічні взаємозв'язки, однак більшість із них не є настільки сильними, щоб свідчити про критичну мультиколінеарність. Це означає, що сформований набір ознак зберігає достатній рівень інформаційної різноманітності та може бути використаний для подальшого кластерного аналізу й побудови економіко-математичних моделей оцінювання сталого розвитку агросектору.

Таблиця 2.4 – Результати ANOVA-аналізу впливу року на обсяги реалізації сільськогосподарської продукції

Продукт	ANOVA p value	Вплив значущий року
Жито	0	ТАК
Птиця свійська	0	ТАК
Свині	0	ТАК
Тварини сільськогосподарські живі	0	ТАК
Культури зернові та зернобобові	0.0002	ТАК
Пшениця	0.004	ТАК
Ячмінь	0.01	ТАК
Картопля	0.01	ТАК
Кукурудза на зерно	0.03	ТАК
Культури овочеві	0.05	Ні
Боби сої	0.14	Ні
Велика рогата худоба	0.5	Ні

Джерело: побудовано автором

Результати однофакторного дисперсійного аналізу, наведені в (табл. 2.4), свідчать про наявність часової неоднорідності у досліджуваному масиві даних, оскільки для більшості видів сільськогосподарської продукції вплив року на обсяг реалізації є статистично значущим. Зокрема, значущий вплив року зафіксовано для жита, птиці свійської, свиней, тварин сільськогосподарських живих, культур

зернових та зернобобових, пшениці, ячменю, картоплі та кукурудзи на зерно, для яких р-значення ANOVA є меншими за традиційний рівень значущості 0,05. Водночас для культур овочевих ($p = 0.05$), бобів сої ($p = 0.14$) та великої рогатої худоби ($p = 0.5$) статистично значущого впливу року не виявлено, що свідчить про відносно стабільніший характер динаміки цих позицій у межах досліджуваного періоду. Це означає, що під час подальшого економіко-математичного моделювання та кластерного аналізу доцільно враховувати наявність міжрічних коливань, адже ігнорування часової мінливості може знизити точність інтерпретації результатів.

Таблиця 2.5 – Результати STL-декомпозиції часових рядів за основними видами сільськогосподарської продукції

Product	Trend Share pct	Seasonality Share pct	Seasonal Strength	Noise Ratio
Боби сої	11.1	97.8	0.98	0.14
Велика рогата худоба	109.5	19.8	0.06	0.48
Жито	64.7	32.7	0.44	0.54
Картопля	51.3	83.4	0.86	0.34
Кукурудза на зерно	59.8	60.7	0.66	0.45
Культури зернові та зернобобові	51.3	68	0.61	0.53
Культури овочеві	124.7	13.6	-0.11	0.52
Птиця свійська	99.2	13.2	-0.02	0.39
Пшениця	27.2	32.1	0.27	0.77
Свині	77.8	52.5	0.81	0.26
Тварини сільськогосподарські живі	49.9	53.7	-0.25	0.83
Ячмінь	20.2	48.4	0.31	0.75

Джерело: побудовано автором

Додатково для поглибленої оцінки часової структури досліджуваних рядів було застосовано STL-декомпозицію, результати якої наведено в (табл. 2.5). Вона дала змогу оцінити співвідношення трендової, сезонної та нерегулярної компонент

у динаміці окремих видів сільськогосподарської продукції. За отриманими результатами найбільш виражена сезонна складова характерна для бобів сої (Seasonality_Share_pct = 97.8, Seasonal_Strength = 0.9) та картоплі (83.4 і 0.8 відповідно), що свідчить про високу залежність цих рядів від сезонних коливань. Відносно сильна сезонність також спостерігається для свиней (52.5, 0.8), кукурудзи на зерно (60.7, 0.6) та культур зернових і зернобобових (68, 0.6). Натомість для великої рогатої худоби, птиці свійської, культур овочевих та тварин сільськогосподарських живих сезонна складова є слабкою або нестійкою, про що свідчать низькі або від'ємні значення показника Seasonal_Strength. Водночас для культур овочевих (Trend_Share_pct = 124.7), великої рогатої худоби (109.5) та птиці свійської (99.2) домінує трендова компонента, що вказує на вищу роль довгострокових змін у формуванні динаміки цих показників. Окремої уваги потребує й нерегулярна складова: найвищий рівень шуму спостерігається для тварин сільськогосподарських живих (Noise_Ratio = 0.83), пшениці (0.7) та ячменю (0.7), що свідчить про більшу нестабільність відповідних рядів. У цілому результати STL-декомпозиції підтверджують, що часові ряди у сформованому датасеті є неоднорідними за своєю структурою, а тому під час подальшого моделювання доцільно враховувати не лише міжрічні відмінності, а й співвідношення трендової, сезонної та випадкової складових.

Узагальнюючи результати проведеного статистичного аналізу, можна зробити висновок, що сформований датасет є придатним для подальшого використання в задачах економіко-математичного моделювання та кластеризації, оскільки охоплює ключові виробничі, ресурсні, інвестиційні та макроекономічні характеристики агровиробництва. Досліджувані змінні виявилися неоднорідними за масштабом вимірювання, рівнем варіативності та характером розподілу. Для частини з них притаманні асиметрія та наявність викидів, що обумовлює необхідність їх попередньої статистичної обробки. Кореляційна структура показників свідчить про відсутність критичної мультиколінеарності, а результати дисперсійного аналізу підтверджують наявність часової неоднорідності для частини досліджуваних видів продукції. Крім того, результати STL-декомпозиції

підтвердили структурну неоднорідність часових рядів за співвідношенням трендової, сезонної та випадкової компонент, що також необхідно враховувати в подальшому моделюванні. У зв'язку з цим перед переходом до наступних етапів аналізу доцільно здійснити стандартизацію змінних та врахувати вплив часового фактора, що підвищить достовірність і аналітичну цінність подальших результатів.

2.2 Побудова та оцінювання моделей прогнозування обсягів агровиробництва

Подальший аналіз присвячено побудові прогнозних моделей обсягів агровиробництва. Прогнозування здійснюється окремо для кожного сегмента аграрного сектору, що дозволяє врахувати специфіку динаміки, рівень сезонності, трендові компоненти та вплив зовнішніх факторів для різних видів сільськогосподарської продукції. Використання кількох моделей дає змогу зіставити їх прогнозну здатність та оцінити доцільність застосування різних підходів для окремих сегментів агровиробництва.

Для оцінювання прогнозної здатності побудованих моделей часові ряди було поділено на навчальну та тестову вибірки у співвідношенні 80/20. Навчальна вибірка використовувалася для побудови моделей, а тестова — для перевірки їх прогнозної здатності. При цьому основна увага приділяється результатам на тестовій вибірці, оскільки саме вони відображають реальну прогнозну здатність моделей у практичних умовах та дозволяють визначити найбільш ефективні підходи для прогнозування обсягів агровиробництва в окремих сегментах.

Сегмент «Боби сої» є одним із важливих напрямів агровиробництва, який характеризується вираженою сезонністю, залежністю від ресурсного забезпечення та чутливістю до цінової й інвестиційної динаміки. Саме тому даний сегмент було обрано для детального аналізу результатів прогнозування, що дозволяє

продемонструвати методику побудови та порівняння різних моделей на конкретному прикладі.

Для прогнозування обсягів реалізації бобів сої за допомогою класичних статистичних методів було застосовано модель ETS (A, N, A), тобто модель з адитивною помилкою, відсутнім трендом та адитивною сезонною складовою. Вибір саме такої специфікації є економічно обґрунтованим, оскільки для даного сегмента характерні регулярні сезонні коливання обсягів реалізації, тоді як довгостроковий тренд у досліджуваному періоді не має стабільно вираженого характеру.

Таблиця 2.6 – Результати оцінювання параметрів моделі ETS(A, N, A) для обсягів реалізації сегмента «Боби сої»

ETS(A, N, A)					
Smoothing parameters					
α		β	γ		φ
0.7425		—	1e-04		—
Initial states					
l			b		
124 844.913			—		
S₀	S₋₁	S₋₂	S₋₃	S₋₄	S₋₅
-160 040	196 224	250 357	30 585	-80 214	-73 495
S₋₆	S₋₇	S₋₈	S₋₉	S₋₁₀	S₋₁₁
-60 419	1134	-2868	33296	53816	-188376
Sigma² = 45 859.53					
AIC		AICc		BIC	
1521.13		1532.29		1552.29	

Джерело: побудовано автором

За результатами оцінювання параметрів експоненційного згладжування коефіцієнт рівня $\alpha = 0,7425$, що свідчить про високу чутливість моделі до нових спостережень та швидку адаптацію рівня ряду до змін у фактичних даних. Водночас параметр сезонності $\gamma = 1 \cdot 10^{-4}$ має дуже мале значення, що означає повільне оновлення сезонних коефіцієнтів і, відповідно, стабільність сезонного профілю бобів сої в межах досліджуваного періоду (табл. 2.6). Початковий рівень ряду оцінено на рівні 124 844,9 т, а початкові сезонні компоненти мають як додатні, так і від’ємні значення, що відображає чергування періодів пікової та низької реалізації продукції протягом року. Дисперсія залишків моделі становить $\sigma^2 =$

45 859,53, а інформаційні критерії $AIC = 1521,13$, $AICc = 1532,29$ та $BIC = 1552,29$ свідчать про прийнятну якість підгонки моделі до даних.

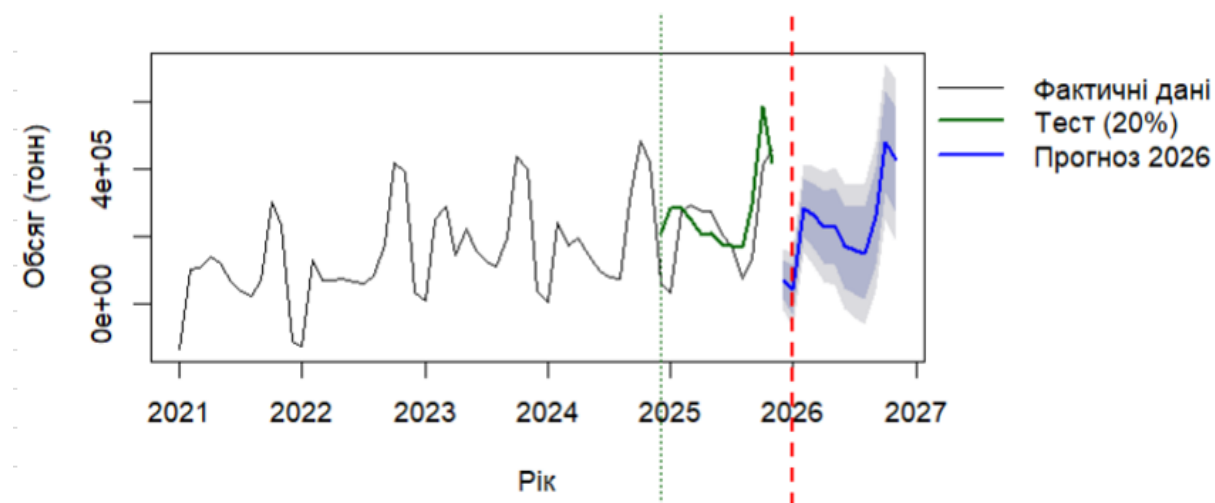


Рисунок. 2. 1 – Прогнозування обсягів реалізації бобів сої за моделлю ETS

Джерело: побудовано автором

Графічні результати прогнозування за моделлю ETS (рис. 2.1) демонструють, що модель загалом коректно відтворює сезонний характер. Аналіз тестової вибірки за 2025 рік показує, що модель досить успішно розпізнає циклічний характер даних. Хоча на початку тестового періоду алгоритм помітно згладжує глибоке просідання, не опускаючись до екстремальних мінімумів фактичних обсягів, отримані результати демонструють здатність моделі коректно відображати динаміку наступного циклу. Зокрема, модель чудово впоралася з прогнозуванням стрімкого пікового зростання наприкінці 2025 року. В прогнозі на 2026 рік очікується традиційне різке падіння на початку періоду, стабілізація на середніх рівнях навесні-влітку та характерний високий пік збору наприкінці року. Важливою деталлю є форма довірчих інтервалів: вони розширюються поступово і залишаються в адекватних межах, що підтверджує загальну стабільність моделі.

У моделі XGBoost для сегмента «Боби сої» як фактори було використано Total_Fert, Avg_Price, Total_Invest та Total_Macro. Такий набір змінних дозволяє врахувати як ресурсні, так і економічні передумови формування обсягів реалізації сої. Додатково в модель було включено лагові значення цільової змінної, що дало змогу поєднати властивості класичного аналізу часових рядів із факторним підходом машинного навчання.

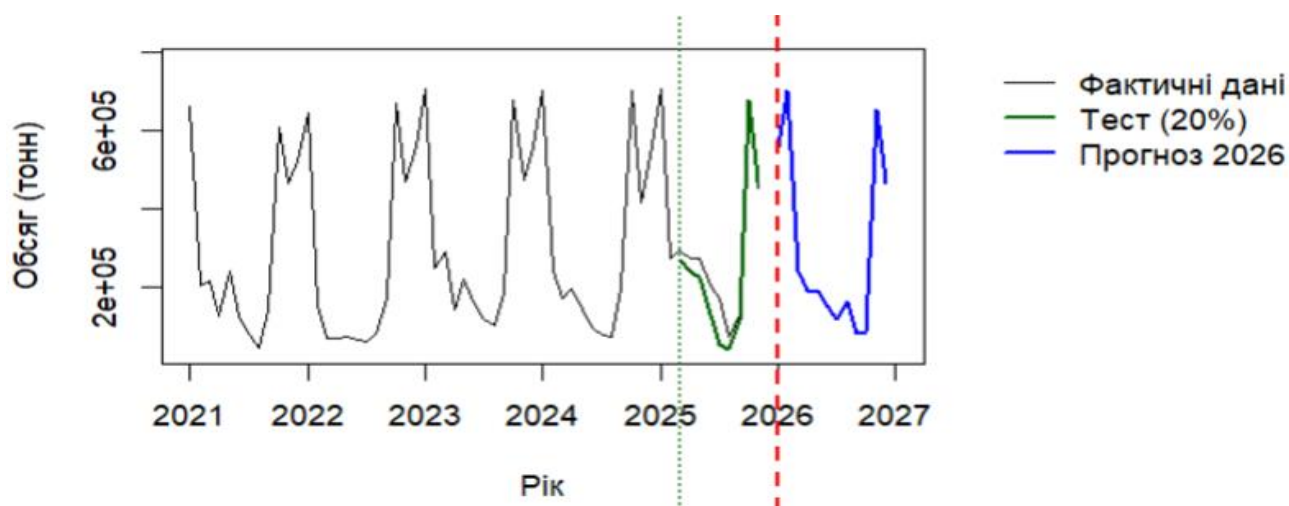


Рисунок. 2. 2 – Прогнозування обсягів реалізації бобів сої за моделлю XGBoost

Джерело: побудовано автором

Результати прогнозування за моделлю XGBoost (рис. 2.2) свідчать про високу точність відтворення фактичної динаміки обсягів реалізації бобів сої. Аналіз тестової вибірки у 2025 році показує, що XGBoost чудово розпізнав циклічну структуру часового ряду. Хоча в період сезонного спаду модель дещо переоцінила глибину просідання, опустившись трохи нижче фактичних значень, вона точно відтворила амплітуду наступного стрімкого зростання наприкінці тестового періоду. В прогнозі на 2026 рік очікується короткочасне утримання високих обсягів на початку року, подальше стрімке зниження з тривалою стабілізацією на мінімальних рівнях протягом весни та літа, після чого прогнозується традиційне різке зростання обсягів наприкінці 2026 року

Згідно з результатами моделі XGBoost, очікуваний обсяг реалізації бобів сої у 2026 році становить 3 597 549 тон, що відповідає зниженню на 10,88 % порівняно з 2025 роком. Отриманий прогноз відображає вплив комплексу факторів, зокрема динаміки ресурсного забезпечення, інвестиційної активності та цінових змін, які не можуть бути враховані моделями експоненційного згладжування.

Наступним етапом аналізу сегмента «Боби сої» стало застосування моделі ARIMAX, яка дозволяє поєднати часову структуру ряду з впливом зовнішніх пояснювальних факторів. Для даного сегмента оптимальною виявилася специфікація $ARIMAX(1,0,0)(1,1,0)[12]$, що включає один авторегресивний

компонент, сезонний авторегресивний елемент із річною періодичністю та сезонне диференціювання. Така структура моделі дає змогу врахувати як інерційність процесу, так і регулярні сезонні коливання обсягів реалізації бобів сої.

Таблиця 2.7 – Оцінені параметри моделі ARIMAX(1,0,0)(1,1,0)[12] для сегмента «Боби сої»

ARIMAX (1,0,0)(1,1,0)[12]							
Coefficients	ar1	sar1	drift	Total_Fert	Avg_Price	Total_Invest	Total_Macro
Estimate	0.51	-0.77	36 20.57	-26.97	-12.49	0.19	1.75
Std. error	0.13	0.08	16 56.48	19.86	12.7	1.16	1.44
Sigma ² = 1.399 × 10 ⁹							
Log likelihood = -563.51							
AIC			AICc			BIC	
1 521.13			1 532.29			1 552.29	

Джерело: побудовано автором

Оцінені параметри моделі, наведені в, свідчать про наявність помірної часової залежності показника: коефіцієнт ar1 = 0,51 вказує на збереження частини динаміки попереднього періоду, тоді як сезонний коефіцієнт sar1 = -0,77 підтверджує суттєвий сезонний ефект із протилежною фазою між відповідними періодами різних років. Додатний параметр drift = 3620,57 відображає середню тенденцію зміни рівня реалізації у часі. Аналіз екзогенних змінних показує, що Total_Fert та Avg_Price мають від'ємні оцінки коефіцієнтів, тоді як Total_Invest і Total_Macro впливають на обсяг реалізації бобів сої позитивно. Така конфігурація коефіцієнтів є економічно обґрунтованою та узгоджується з логікою впливу інвестиційної активності й загального стану аграрного сектору на обсяги реалізації. Якість підгонки моделі ARIMAX характеризується значеннями AIC = 1521,13, AICc = 1532,29 та BIC = 1552,29, а також оцінкою дисперсії залишків $\sigma^2 = 1,399 \times 10^9$, що свідчить про прийнятний рівень узгодження моделі з даними (табл. 2.7).

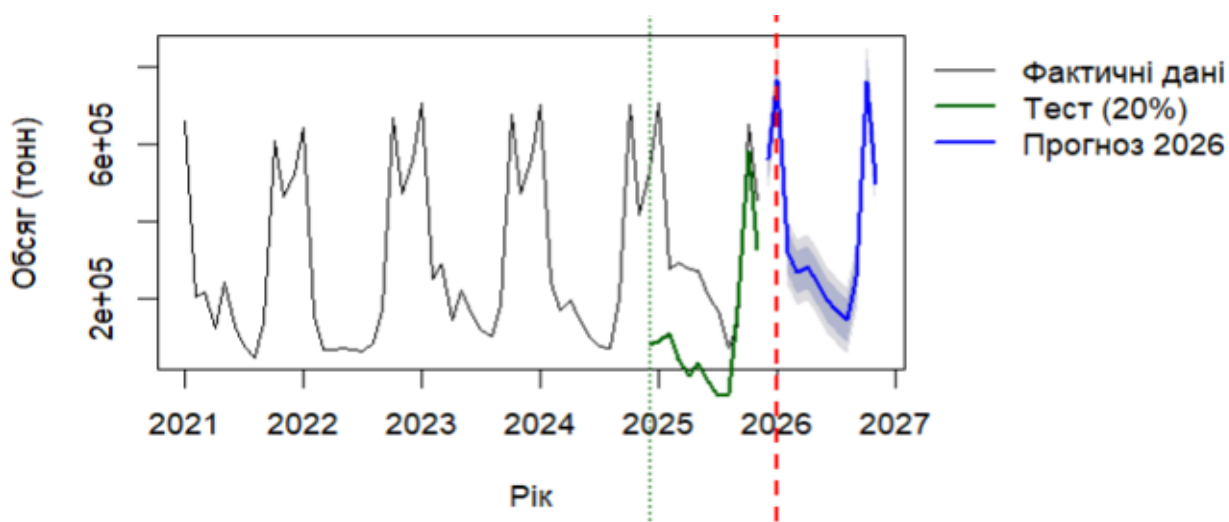


Рисунок. 2.3 – Прогнозування обсягів реалізації бобів сої за моделлю ARIMAX

Джерело: побудовано автором

Графічне представлення результатів прогнозування (рис. 2.3) показує, що на тестовій вибірці, що охоплює 2025 рік, вказує на те, що модель частково вловлює циклічну природу даних, але має проблеми з калібруванням амплітуди в періоди спаду. З одного боку, ARIMAX досить успішно ідентифікувала час початку та загальний напрям пікового зростання наприкінці року. З іншого боку, у міжсезоння (перша половина 2025 року) модель продемонструвала гіпертрофовану реакцію на спад: зелена лінія просідає значно глибше за фактичні показники, опускаючись помітно нижче традиційних історичних мінімумів. Це свідчить про те, що вплив додаткових факторів у моделі може бути надмірно обтяжливим і спотворювати природне «дно» циклу. Прогноз на 2026 рік загалом зберігає притаманну ряду загальну структуру.

Для врахування можливих нелінійних залежностей у динаміці обсягів реалізації бобів сої в межах даного сегмента додатково було застосовано нейромережеву модель NNETAR, яка належить до класу штучних нейронних мереж для часових рядів. Для досліджуваного показника оптимальною виявилася модель NNETAR(2,1,4)[12], що використовує два лаги авторегресивної частини, одне диференціювання та сезонну компоненту з періодичністю 12. Навчання здійснювалося шляхом усереднення результатів 20 нейронних мереж, кожна з яких

має архітектуру 7–4–1 та містить 37 вагових коефіцієнтів, що дозволяє зменшити вплив випадковості початкових умов і підвищити стабільність прогнозу.

Таблиця 2.8 – Параметри та характеристика моделі NNETAR(2,1,4)[12] для сегмента «Боби сої»

NNAR(2, 1, 4)[12]
Average of 20 networks, each of which is a 7–4–1 network with 37 weights
options were – linear output units
$\sigma^2 = 81\,330\,177$

Джерело: побудовано автором

Оцінка якості моделі NNETAR, наведена в (табл. 2.8), показує, що на навчальній вибірці нейромережевий підхід демонструє відносно прийнятні значення похибок, що свідчить про здатність моделі відтворювати складну структуру часового ряду. Водночас дисперсія залишків $\sigma^2 = 81\,330\,177$ є суттєво вищою порівняно з класичною моделлю ETS та економетричною моделлю ARIMAX, що вказує на більшу нестабільність прогнозів.

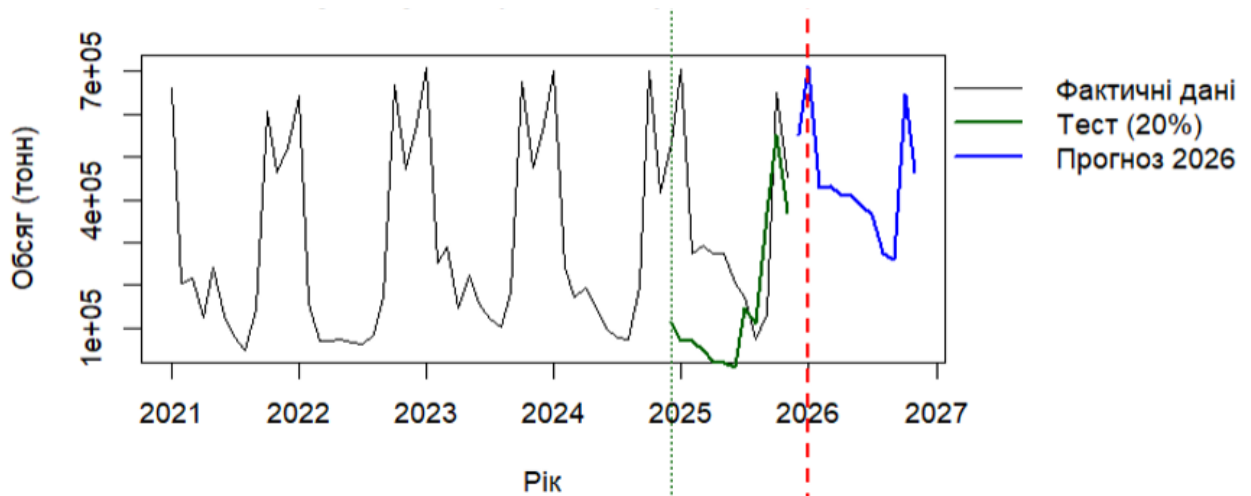


Рисунок. 2. 4 – Прогнозування обсягів реалізації бобів сої за моделлю NNETAR

Джерело: побудовано автором

Графічне зіставлення фактичних і прогнозних значень (рис. 2.4) підтверджує отримані кількісні результати. Прогноз, сформований за допомогою NNETAR, загалом відтворює сезонну структуру динаміки бобів сої, однак на початку тестового періоду прогноз різко обвалюється, опускаючись значно нижче фактичних значень і досягаючи неприродного низького рівня, хоча модель все ж

таки спромоглася вийти на загальний рівень високих обсягів наприкінці року. Прогноз на 2026 рік виглядає специфічно. Замість типового глибокого U-подібного або V-подібного спаду, характерного для міжсезоння 2021–2024 років, модель показує траєкторію у вигляді не різкого падіння, а поступового зниження після якого настає миттєвий вертикальний стрибок наприкінці року.

Таблиця 2.9 – Порівняння точності прогнозних моделей для сегмента «Боби сої» на навчальній та тестовій вибірках

Training set	ME	RMSE	MAE	MPE	MAPE	MASE
XGBOOST	84.4	1 133.3	843.4	-0.1	0.5	0.005
ETS	2 410.6	40 050.7	27 853.4	-1.9	23.7	0.4
ARIMAX	-496.7	30 800.5	20031.2	-4.38	13.73	0.35
NNETAR	-11.4	9 018.3	6 690.9	-1.4	4.8	0.1

Testing set	ME	RMSE	MAE	MPE	MAPE	MASE
XGBOOST	34 129.7	52 598	42 539.4	21.5	23	0.3
ETS	-31 832	117 209	80 511.9	-27.1	44.3	1.9
ARIMAX	127 559	210 553	196 652	62.3	90.5	2.9
NNETAR	75 379.9	209 367	187 662	15.6	83.9	2.8

Джерело: побудовано автором

Після детального розгляду результатів прогнозування обсягів реалізації бобів сої за допомогою моделей ETS, XGBoost, ARIMAX та NNETAR, доцільно здійснити їх узагальнене порівняння за показниками точності. Відповідні результати оцінювання на навчальній та тестовій вибірках наведено в (табл. 2.9), що дозволяє комплексно оцінити як здатність моделей до відтворення історичної динаміки, так і їхню реальну прогнозну спроможність.

Аналіз результатів на навчальній вибірці показує суттєву перевагу моделі XGBoost, яка демонструє найменші значення ключових показників похибок: RMSE = 1133,3, MAE = 843,4, MAPE = 0,5 % та MASE = 0,005. Це свідчить про високу здатність моделі точно відтворювати структуру часового ряду. Модель NNETAR також характеризується відносно низькими похибками (RMSE = 9018,3, MAE = 6690,9), однак істотно поступається XGBoost за всіма критеріями. Натомість класичні підходи ETS та ARIMAX демонструють значно вищі значення помилок,

що вказує на обмежені можливості цих моделей у відтворенні складної динаміки обсягів реалізації бобів сої навіть на етапі навчання.

Найбільш показовими є результати на тестовій вибірці, які відображають реальну прогностну здатність моделей. Як видно з (табл. 2.9), модель XGBoost зберігає найкращі показники точності і на етапі перевірки ($RMSE = 52\,597,97$, $MAE = 42\,539,4$, $MAPE = 23\%$, $MASE = 0,27$), що свідчить про її високу стійкість до структурних змін у даних. Моделі ETS та ARIMAX демонструють різке зростання похибок на тестовій вибірці, зокрема значення MAPE перевищують 40–90 %, що вказує на низьку адаптивність цих підходів до змін у кінцевій частині ряду. Модель NNETAR, хоча й перевершує ARIMAX за окремими показниками, також характеризується значною нестабільністю прогнозів ($MAPE = 83,9\%$, $MASE = 2,8$), що обмежує її практичну придатність для прогнозування в даному сегменті.

Таким чином, результати порівняльного аналізу підтверджують, що для сегмента «Боби сої» модель XGBoost є найбільш ефективною як з точки зору точності прогнозу, так і з погляду стійкості до змін у часовій динаміці. Класичні статистичні та економетричні моделі (ETS і ARIMAX) доцільно використовувати як базові або пояснювальні інструменти, тоді як нейромережева модель NNETAR може розглядатися як допоміжний підхід для виявлення нелінійних ефектів. Отримані висновки щодо сегмента «Боби сої» слугують методичною основою для узагальненого аналізу результатів прогнозування за іншими сегментами агровиробництва, який наведено у наступному підрозділі.

Сегмент «Велика рогата худоба» належить до ключових складових тваринницької галузі та характеризується відносно стабільною динамікою обсягів реалізації, менш вираженою сезонністю порівняно з рослинницькими культурами та значною залежністю від довгострокових виробничих і економічних факторів. На відміну від сегмента «Боби сої», для якого характерні різкі сезонні коливання, динаміка реалізації продукції великої рогатої худоби є більш згладженою, що обумовлює доцільність використання як класичних статистичних моделей, так і факторних та машинно-навчальних підходів.

Для прогнозування обсягів реалізації сегмента «Велика рогата худоба» базово застосовано модель ETS., яка використовується як базовий статистичний інструмент для опису часових рядів зі стабільною динамікою. За результатами автоматичного відбору оптимальною виявилася модель ETS(A, N, N), тобто модель з адитивною помилкою, без трендової та сезонної складових. Така специфікація є логічною з огляду на характер досліджуваного ряду, для якого не спостерігається чітко вираженої сезонності або стійкого тренду в межах аналізованого періоду.

Таблиця 2.10 – Результати оцінювання параметрів моделі ETS(A, N, N) для обсягів реалізації сегмента «Велика рогата худоба»

ETS(A, N, N)					
Smoothing parameters					
α		β	γ		ϕ
0.7286		—	—		—
Initial states					
l			b		
9878.6449			—		
S₀	S₋₁	S₋₂	S₋₃	S₋₄	S₋₅
—	—	—	—	—	—
S₋₆	S₋₇	S₋₈	S₋₉	S₋₁₀	S₋₁₁
—	—	—	—	—	—
Sigma² = 1 115.368					
AIC		AICc		BIC	
1 072.54		1 072.98		1 078.77	

Джерело: побудовано автором

Оцінені параметри моделі, наведені в (табл. 2.10), свідчать що коефіцієнт згладжування рівня $\alpha = 0,7286$, що свідчить про достатньо швидке реагування моделі на нові спостереження. Початковий рівень ряду оцінено на рівні 9 878,64 т, що відповідає середньому масштабу реалізації продукції великої рогатої худоби у вибірці. Відсутність параметрів тренду та сезонності підтверджує припущення про відносну рівномірність динаміки даного сегмента. Дисперсія залишків моделі становить $\sigma^2 = 1 115,37$, а значення інформаційних критеріїв AIC = 1072,54, AICc = 1072,98 та BIC = 1078,77 свідчать про добру якість підгонки моделі до фактичних даних.

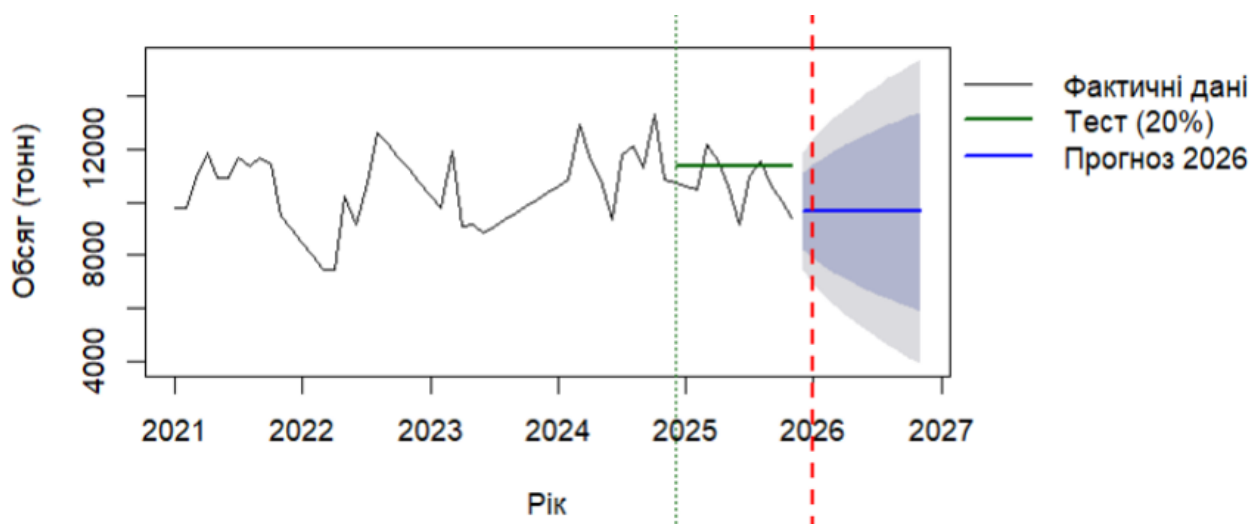


Рисунок. 2. 5 – Прогнозування обсягів реалізації великої рогатої худоби за моделлю ETS

Джерело: побудовано автором

Графічне відображення результатів прогнозування (рис. 2.5) демонструє, в тестовій вибірці, що охоплює 2025 рік, повну нездатність моделі відтворити фактичну мінливість часового ряду. Замість того, щоб слідувати за реальними коливаннями, алгоритм згенерував пласку, горизонтальну траєкторію.

Прогноз на 2026 рік є прямим продовженням цієї математичної логіки. Модель генерує стабільну пряму лінію, не передбачаючи жодних змін, циклів чи структурних зрушень. При цьому критично важливою візуальною деталлю є форма довірчих інтервалів. Вони стрімко розходяться, що є індикатором експоненційного зростання математичної невизначеності з кожним наступним кроком багатокрокового прогнозування.

Подальший аналіз сегмента було здійснено із застосуванням моделі XGBoost з використанням факторів: Total_Animals, Avg_Price, Total_Invest та Total_Macro. Обраний набір змінних дозволяє комплексно врахувати як виробничу основу сегмента, так і макроекономічні умови його функціонування.

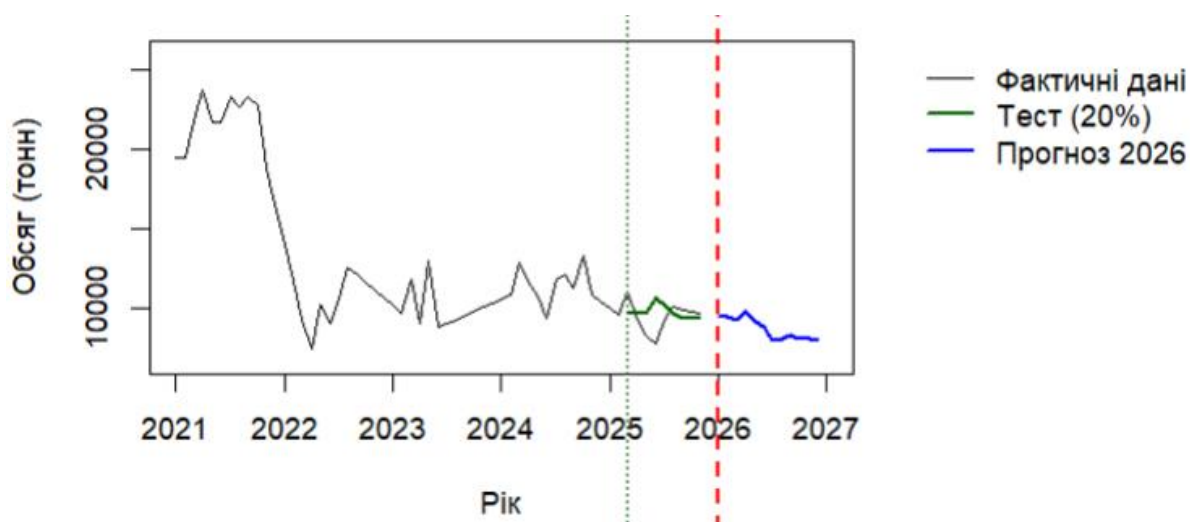


Рисунок. 2. 6 – Прогнозування обсягів реалізації великої рогатої худоби за моделлю XGBoost

Джерело: побудовано автором

Графічне зіставлення фактичних і прогнозних значень (рис. 2.6) підтверджує отримані кількісні результати. Прогноз XGBoost для сегмента «Велика рогата худоба» характеризується плавною, але адаптивною траєкторією, яка коректно відображає зниження рівня реалізації в кінцевій частині ряду. Згідно з результатами моделі, очікуваний обсяг реалізації у 2026 році становить 105 141,6 т, що відповідає зменшенню приблизно на 8,95 % порівняно з 2025 роком. Такий прогноз узгоджується з попередньо виявленою тенденцією до зниження реалізації у цьому сегменті та відображає вплив змін у поголів'ї, ціновій динаміці й інвестиційній активності.

Подальший етап аналізу сегмента «Велика рогата худоба» передбачав застосування моделі ARIMAX. Для даного сегмента оптимальною була визначена модель $ARIMAX(1,0,0)(0,0,1)[12]$, що включає один авторегресивний компонент та сезонний ковзний середній елемент з річною періодичністю. Така специфікація відображає інерційний характер динаміки реалізації продукції великої рогатої худоби та наявність помірних сезонних коливань без необхідності сезонного диференціювання.

Таблиця 2.11 – Оцінені параметри моделі ARIMAX(1,0,0)(0,0,1)[12] для сегмента «Велика рогата худоба»

ARIMAX (1,0,0)(0,0,1)[12]						
Coefficients	ar1	sma1	Total_Animals	Avg_Price	Total_Invest	Total_Macro
Estimate	0.91	0.29	0	-0.11	0.01	2.05
Std. error	0.05	0.18	0.0002	0.11	0.11	1.09
Sigma ² = 2 584 367						
Log likelihood = -517.63						
AIC			AICc		BIC	
1 049.27			1 051.46		1 063.81	

Джерело: побудовано автором

Оцінені параметри моделі, наведені в (табл. 2.11), свідчать про високу залежність поточного рівня реалізації від попередніх значень ряду: коефіцієнт $ar1 = 0,9123$ вказує на значну інерційність процесу. Сезонний компонент $sma1 = 0,2903$ підтверджує наявність сезонного ефекту, однак його вплив є помірним, що узгоджується з відносно стабільною динамікою тваринницького сегмента. Аналіз екзогенних змінних показує, що кількість *Total_Animals* не чинить істотного впливу на короткострокову динаміку реалізації, тоді як *Avg_Price* має від'ємний коефіцієнт, що може свідчити про стримування обсягів збуту за умов зростання цін. Водночас *Total_Invest* та *Total_Macro* впливають на обсяги реалізації позитивно, відображаючи роль довгострокових економічних чинників у розвитку галузі.

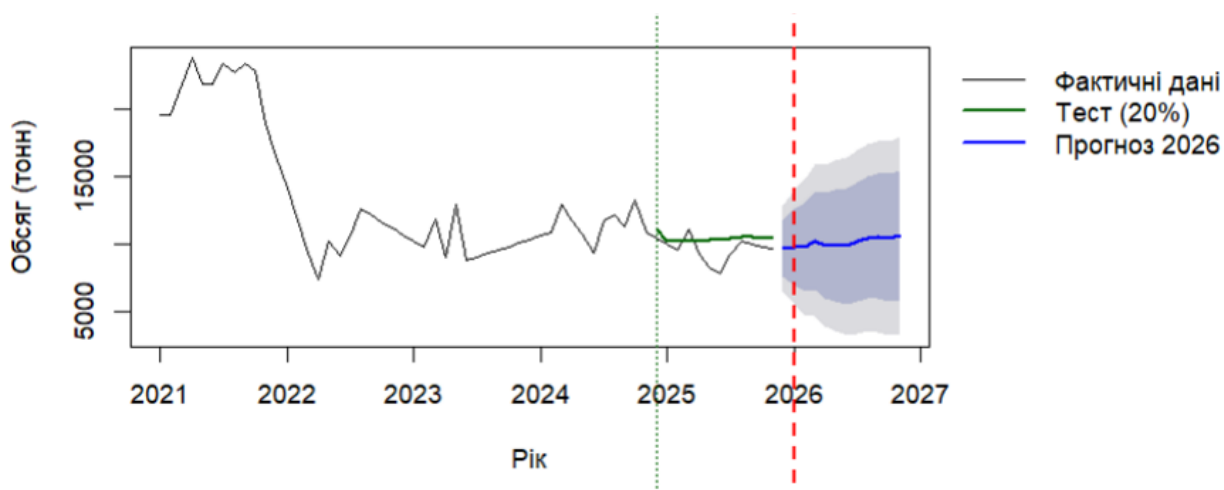


Рисунок. 2. 7 – Прогнозування обсягів реалізації великої рогатої худоби за моделлю ARIMAX

Джерело: побудовано автором

Якість підгонки моделі ARIMAX характеризується значеннями $AIC = 1049,27$, $AICc = 1051,46$ та $BIC = 1063,81$, а також оцінкою дисперсії залишків $\sigma^2 = 2\,584\,367$, що свідчить про задовільний рівень узгодження моделі з фактичними даними (табл. 2.11). Графічне представлення результатів прогнозування (рис. 2.7) показує, в тестовій вибірці, що припадає на 2025 рік, фактичну нездатність моделі відтворити локальну мінливість часового ряду. Замість того, щоб реагувати на коливання фактичних даних, тестовий прогноз формує практично ідеальну пряму горизонтальну лінію. Прогноз на 2026 рік є логічним продовженням цієї інерційної тенденції. Модель малює стабільну траєкторію з ледь помітними, флуктуаціями. Однак головним аналітичним маркером тут є не сама лінія прогнозу, а надзвичайно широкі довірчі інтервали. Їхнє стрімке розширення беззаперечно свідчить про те, що за видимою стабільністю точкового прогнозу ховається математична невизначеність.

На завершальному етапі аналізу сегмента для перевірки можливості врахування нелінійних закономірностей у динаміці обсягів реалізації було застосовано модель NNETAR. Для досліджуваного часового ряду оптимальною виявилася специфікація $NNETAR(4,1,5)[12]$, яка включає чотири лаги авторегресивної частини, одне диференціювання та сезонну компоненту з періодичністю 12. Навчання моделі здійснювалося шляхом усереднення результатів 20 нейронних мереж, кожна з яких має архітектуру 9–5–1 та містить 56 вагових коефіцієнтів, що дозволяє зменшити вплив випадковості початкових умов і підвищити стабільність прогнозу.

Таблиця 2.12 – Параметри та характеристика моделі $NNETAR(4,1,5)[12]$ для сегмента «Велика рогата худоба»

$NNAR(4, 1, 5)[12]$
Average of 20 networks, each of which is a 9–5–1 network with 56 weights
options were – linear output units
$\sigma^2 = 9\,803$

Джерело: побудовано автором

Показники якості прогнозування моделі NNETAR, наведені в (табл. 2.12), свідчать, що модель достатньо добре відтворює загальну структуру часового ряду

на навчальній вибірці. Оцінка дисперсії залишків $\sigma^2 = 9\,803$ вказує на прийнятний рівень внутрішньої узгодженості моделі. Водночас на тестовій вибірці спостерігається зростання похибок прогнозування, що свідчить про обмежену здатність моделі до узагальнення за умов змін у кінцевій частині ряду.

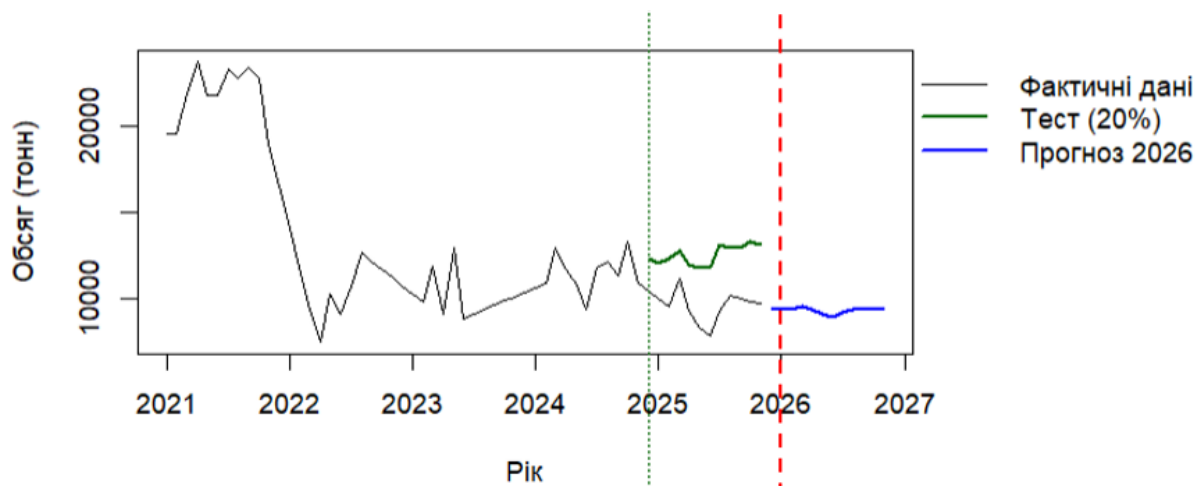


Рисунок. 2. 8 – Прогнозування обсягів реалізації великої рогатої худоби за моделлю NNETAR

Джерело: побудовано автором

Графічне зіставлення фактичних і прогнозних значень (рис. 2.8) підтверджує отримані кількісні результати. Аналіз тестової вибірки, яка охоплює 2025 рік, чітко демонструє розсинхронізацію моделі з реальними ринковими даними. У той час як фактичні обсяги продовжують демонструвати спадну, динаміку з оновленням локальних мінімумів, тестовий прогноз NNETAR формує завищену, дещо кутасту та ступінчасту траєкторію. Модель систематично переоцінює обсяги, проходячи помітно вище реальних значень, що свідчить про її нездатність адаптуватися до тривалого депресивного тренду в даних. Прогноз на 2026 рік демонструє ще більш аномальну та суперечливу поведінку. Спостерігається різкий, логічно не обґрунтований розрив рівнів: після завищених показників наприкінці тестового періоду прогноз стрімко «обвалюється» до позначки близько 10 000 тонн і перетворюється на практично пласку, згладжену лінію з мінімальними флуктуаціями. Така неприродна стабільність та втрата амплітуди не відповідають історичному характеру ряду.

Після детального розгляду результатів прогнозування обсягів реалізації продукції великої рогатої худоби за допомогою моделей ETS, XGBoost, ARIMAX та NNETAR доцільно здійснити їх узагальнене порівняння за показниками точності на навчальній і тестовій вибірках. Відповідні результати оцінювання наведено в (табл. 2.13), що дозволяє зіставити здатність моделей як до відтворення історичної динаміки, так і до формування надійного прогнозу поза межами періоду навчання.

Таблиця 2.13 – Порівняння точності прогнозних моделей для сегмента «Велика рогата худоба» на навчальній та тестовій вибірках

Training set	ME	RMSE	MAE	MPE	MAPE	MASE
XGBOOST	3.6	70.4	54.9	-0.01	0.5	0.04
ETS	-5.4	1 096.3	862.9	-0.7	8.2	0.6
ARIMAX	-143	1 523.7	1 161	-2.6	9.9	0.3
NNETAR	-1.65	99	68.5	-0.06	0.7	0.02

Testing set	ME	RMSE	MAE	MPE	MAPE	MASE
XGBOOST	-279	1236	977.7	-4.2	10.9	0.8
ETS	70 800	211 427.2	71 289.5	13.05	17.8	31.9
ARIMAX	71 801	211 710.1	71 928.2	20.8	22.1	14.9
NNETAR	69 995	211 008.2	70 991	7.4	16.93	14.7

Джерело: побудовано автором

Аналіз результатів на навчальній вибірці показує, що найкращі значення показників похибок демонструє модель XGBoost, для якої зафіксовано найменші значення $RMSE = 70,4$, $MAE = 54,9$ та $MAPE = 0,5\%$. Це свідчить про високу здатність моделі відтворювати внутрішню структуру часового ряду сегмента «Велика рогата худоба». Модель NNETAR також демонструє відносно добрі результати на навчальній вибірці ($RMSE = 99$, $MAE = 68,5$), однак поступається XGBoost за всіма ключовими метриками. Натомість моделі ETS та ARIMAX характеризуються значно вищими значеннями похибок, що вказує на обмежену здатність цих моделей точно відтворювати динаміку навіть у межах навчального періоду.

Найбільш показовими є результати на тестовій вибірці, які відображають реальну прогнозну спроможність моделей. Як видно з (табл. 2.13), модель XGBoost зберігає найкращі показники і на етапі перевірки ($RMSE = 1\,236$, $MAE = 977,7$, $MAPE = 10,9\%$, $MASE = 0,8$), що свідчить про її високу стійкість до змін у кінцевій

частині ряду. Моделі ETS, ARIMAX та NNETAR демонструють різке зростання похибок на тестовій вибірці, зокрема значення RMSE перевищують 200 тис. тонн, а MAPE знаходиться в межах 16–22 % і більше, що вказує на суттєве погіршення прогнозової якості. Це свідчить про обмежену здатність цих моделей до узагальнення в умовах структурних змін і нестабільності динаміки.

Результати порівняльного аналізу підтверджують, що для сегмента «Велика рогата худоба» модель XGBoost є найбільш ефективною з точки зору точності та стабільності прогнозу як на навчальній, так і на тестовій вибірках. Модель NNETAR може розглядатися як допоміжний інструмент для аналізу нелінійних ефектів, тоді як ETS та ARIMAX доцільно використовувати як базові або пояснювальні підходи.

Сегмент «Кукурудза на зерно» є одним із найбільш масштабних і економічно значущих напрямів рослинницького виробництва, який характеризується високими обсягами реалізації, чітко вираженою сезонністю та значною залежністю від погодних, технологічних і ринкових факторів.

Таблиця 2.14 – Результати оцінювання параметрів моделі ETS(A, N, A) для обсягів реалізації «Кукурудза на зерно»

ETS(A, N, A)					
Smoothing parameters					
α		β	γ		ϕ
0.99		—	1e-04		—
Initial states					
I			b		
1 429 749.02			—		
S₀	S₋₁	S₋₂	S₋₃	S₋₄	S₋₅
550 223.8	642 159.1	363 004	−579 416.3	−784 334.5	−782 389.7
S₋₆	S₋₇	S₋₈	S₋₉	S₋₁₀	S₋₁₁
−283 705	−187 464.5	−41 011.79	283 085.7	363 039.1	456 810.1
Sigma² = 411 216.2					
AIC		AICc		BIC	
1779.96		1791.13		1811.13	

Джерело: побудовано автором

Для прогнозування обсягів реалізації кукурудзи на зерно базово застосовано модель ETS. За результатами автоматичного відбору оптимальною для даного

сегмента виявилася модель ETS(A, N, A), тобто модель з адитивною помилкою, відсутнім трендом та адитивною сезонною складовою. Така специфікація є обґрунтованою, оскільки сезонні коливання реалізації кукурудзи мають стабільний характер, тоді як довгостроковий тренд у межах досліджуваного періоду не є стійко вираженим.

Оцінені параметри моделі, наведені в (табл. 2.14), свідчать про надзвичайно високу чутливість до нових спостережень: коефіцієнт згладжування рівня $\alpha = 0,9999$, що означає практично миттєву адаптацію рівня ряду до змін у фактичних даних. Параметр сезонності $\gamma = 1 \cdot 10^{-4}$ має дуже мале значення, що вказує на стабільність сезонного профілю реалізації кукурудзи на зерно в межах аналізованого періоду. Початковий рівень ряду оцінено на рівні 1 429 749,0 т, що відображає масштабність цього сегмента. Початкові сезонні компоненти мають значні додатні й від'ємні значення, що свідчить про різко виражені пікові та мінімальні періоди реалізації протягом року. Дисперсія залишків моделі становить $\sigma^2 = 411\,216,2$, а значення інформаційних критеріїв AIC = 1779,96, AICc = 1791,13 та BIC = 1811,13 свідчать про прийнятну якість підгонки моделі до фактичних даних.

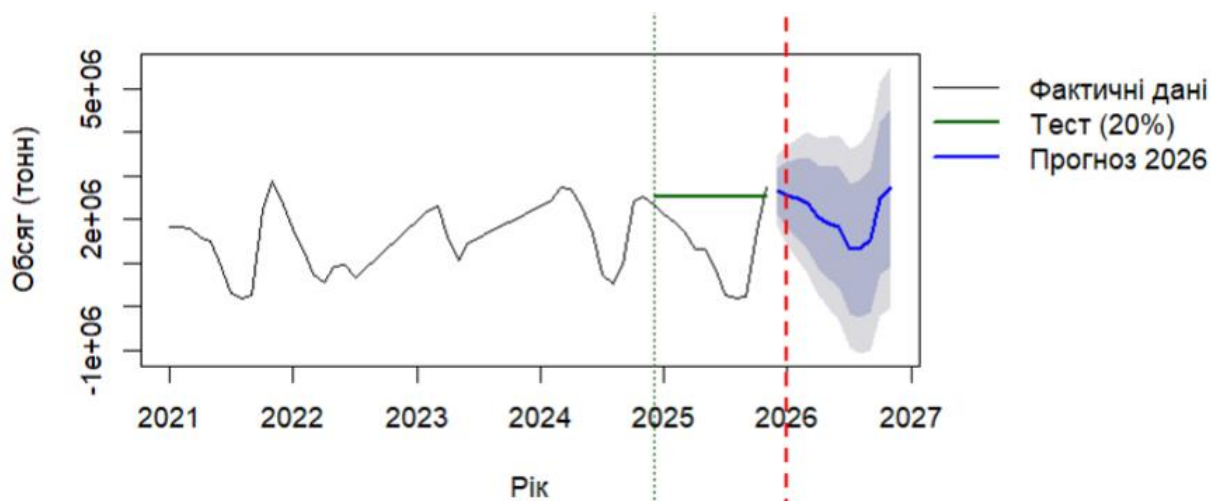


Рисунок. 2. 9 – Прогнозування обсягів реалізації кукурудзи на зерно за моделлю ETS

Джерело: побудовано автором

Графічне представлення результатів прогнозування (рис. 2.9) демонструє, що тестова частина моделі суттєво відрізняється від фактичної динаміки ряду. У

той час як реальні значення у 2025 році демонструють спад та різкі зміни показника, модель ETS формує майже горизонтальну траєкторію без помітного реагування на коливання. Це свідчить про недостатню здатність моделі точно відтворювати короткострокові зміни та високу волатильність часового ряду.

Прогноз на 2026 рік показує хвилюподібну динаміку: на початку року очікується поступове зниження обсягів реалізації, після чого наприкінці прогностного періоду передбачається певне відновлення показника. Водночас довірчі інтервали прогнозу є досить широкими та розширюються з часом, що свідчить про високий рівень невизначеності прогностних оцінок. Це вказує на обмежену точність моделі ETS для прогнозування реалізації кукурудзи на зерно в умовах нестабільної динаміки ряду.

Подальший аналіз сегмента було здійснено із застосуванням моделі XGBoost, як пояснювальні змінні було використано Total_Area, Total_Fert, Avg_Price, Total_Invest та Total_Macro.

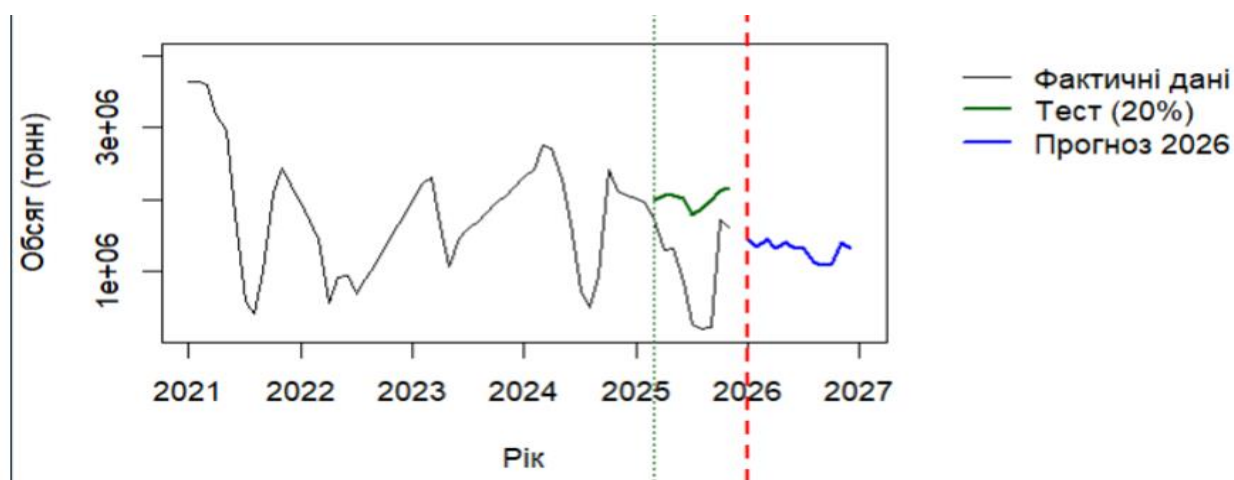


Рисунок. 2. 10 – Прогнозування обсягів реалізації кукурудзи на зерно за моделлю XGBoost

Джерело: побудовано автором

Графічне зіставлення фактичних і прогностних значень (рис. 2.10) свідчить про те, що модель XGBoost для обсягів кукурудзи чітко показує наявність суттєвих відхилень між реальними та прогнозованими показниками на етапі тестування. Аналіз тестової вибірки, яка починається у 2025 році, показує, що модель абсолютно не вловила різкого спаду обсягів. У той час як фактичний ряд

демонструє глибоке просідання з наступним відновленням, тестовий прогноз залишається на стабільно високому рівні, майже не реагуючи на ці зміни. Це свідчить про те, що модель XGBoost у даній конфігурації згладжує результати та має вкрай обмежену здатність адаптуватися до високої волатильності даних та глибоких сезонних (або аномальних) коливань.

Прогноз на 2026 рік також демонструє відносно згладжену траєкторію, яка коливається на рівні, значно нижчому за історичні максимуми попередніх років (т. Очікується стабільна динаміка з незначними флуктуаціями, без різких стрибків чи падінь, що були характерні для періоду 2021–2024 років. Водночас, з огляду на очевидну неспроможність моделі відтворити амплітуду коливань на тестовому періоді 2025 року, отриманий прогноз на 2026 рік слід інтерпретувати з великою обережністю. Згідно з результатами моделі XGBoost, очікуваний обсяг реалізації кукурудзи на зерно у 2026 році становить 15 680 056 т, що відповідає зростанню приблизно на 2,69 % порівняно з 2025 роком.

Таблиця 2.15 – Оцінені параметри моделі ARIMAX(1,0,0)(1,1,0)[12] для сегмента «Кукурудза на зерно»

ARIMAX (1,0,1)(0,1,1)[12]									
Coefficients	ar1	ma1	sma1	drift	Total Area	Total Fert	Avg Price	Total Invest	Total Macro
Estimate	0.4	0.6	−0.7	76 387	4.1	−1 500.6	−966.7	−56.6	−430.2
Std. error	0.2	0.1	0.4	24 858.7	1.5	610.3	351.1	17	175.5
Sigma ² = 1.199 × 10 ¹¹									
Log likelihood = −665.8									
AIC				AICc			BIC		
1351.57				1357.68			1370.07		

Джерело: побудовано автором

Подальший аналіз сегмента було продовжено із застосуванням моделі ARIMAX. Для даного сегмента оптимальною виявилася специфікація ARIMAX(1,0,1)(0,1,1)[12], що включає авторегресивний та ковзний середній компоненти, сезонний компонент ковзного середнього з річною періодичністю, а також сезонне диференціювання.

Оцінені параметри моделі, наведені в (табл. 2.15), свідчать про наявність як часової інерції, так і суттєвого впливу зовнішніх факторів. Додатний коефіцієнт ar1

$= 0,4$ вказує на помірну залежність поточного рівня реалізації від попередніх значень, тоді як коефіцієнт $ma1 = 0,6$ відображає роль короткострокових випадкових шоків. Від’ємний сезонний компонент $sma1 = -0,7$ підтверджує суттєву сезонну асиметрію між відповідними періодами різних років. Наявність додатного параметра $drift = 76\,387$ свідчить про загальну тенденцію до зростання базового рівня реалізації в часі.

Аналіз екзогенних змінних показує, що $Total_Area$ чинить істотний додатний вплив на обсяги реалізації кукурудзи на зерно, що є економічно логічним для культури з високою залежністю від посівних площ. Водночас $Total_Fert$ та Avg_Price мають від’ємні коефіцієнти, що може свідчити про зниження фізичних обсягів реалізації в умовах інтенсифікації виробництва або зростання цін. Вплив $Total_Invest$ та $Total_Macro$ також має від’ємний характер, що може відображати перерозподіл ресурсів між культурами в роки високої аграрної активності. Якість підгонки моделі $AIC = 1351,57$, $AICc = 1357,68$ та $BIC = 1370,07$, а також оцінкою дисперсії залишків $\sigma^2 = 1,199 \times 10^{11}$, що свідчить про прийнятний рівень узгодження моделі з фактичними даними (табл. 2.15).

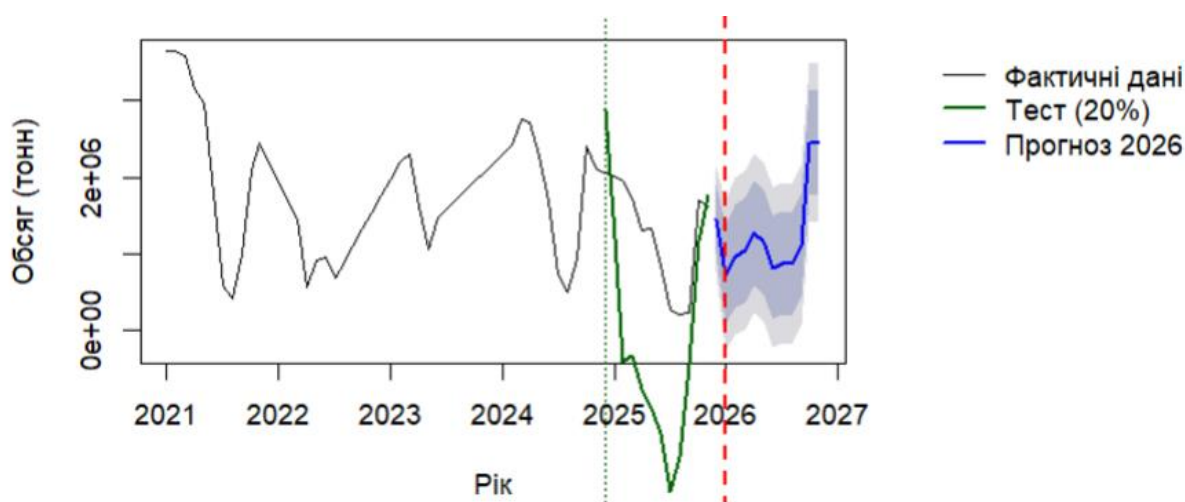


Рисунок. 2. 11 – Прогнозування обсягів реалізації кукурудзи на зерно за моделлю ARIMAX

Джерело: побудовано автором

Графічне зіставлення фактичних і прогнозних значень багаторфакторної моделі ARIMAX на (рис.2.11) свідчить про її надзвичайно високу чутливість до змін у динаміці часового ряду. Аналіз тестової вибірки, яка охоплює 2025 рік,

показує, що на відміну від згладжування, модель намагається відтворити різке падіння обсягів, однак робить це з надмірною, амплітудою. Прогнозні значення на тестовому проміжку просідають значно глибше за фактичні показники, опускаючись до екстремальних мінімумів, після чого відбувається такий самий стрімкий відскок вгору. Це вказує на те, що модель переоцінює глибину кризових коливань, ймовірно, через специфічну реакцію на включені зовнішні фактори або структурні зміни. Прогноз на 2026 рік демонструє досить нестабільну траєкторію з нетиповою динамікою. Після помірних коливань у першій половині року очікується різке зростання обсягів наприкінці періоду. Важливою візуальною деталлю є наявність широких довірчих інтервалів, що безпосередньо вказує на високий рівень статистичної невизначеності та широкі межі можливої похибки цього прогнозу.

На завершальному етапі аналізу сегмента було застосовано модель NNETAR. Для досліджуваного часового ряду оптимальною виявилася специфікація NNETAR(2,1,4)[12], яка включає два лаги авторегресивної частини, одне диференціювання та сезонну компоненту з періодичністю 12. Навчання моделі здійснювалося шляхом усереднення результатів 20 нейронних мереж, кожна з яких має архітектуру 8–4–1 та містить 41 ваговий коефіцієнт, що дозволяє зменшити вплив випадкових початкових умов і підвищити стабільність прогнозних оцінок.

Таблиця 2. 16 Параметри та характеристика моделі NNETAR(4,1,5)[12] для сегмента «Кукурудза на зерно»

NNAR(2, 1, 4)[12]
Average of 20 networks, each of which is an 8–4–1 network with 41 weights
options were – linear output units
$\sigma^2 = 2.491 \times 10^9$

Джерело: побудовано автором

Показники якості прогнозування моделі NNETAR, наведені в (табл. 2.16), свідчать, що нейромережева модель здатна відтворювати загальну форму динаміки обсягів реалізації кукурудзи на зерно на навчальній вибірці. Водночас оцінка дисперсії залишків $\sigma^2 = 2,491 \times 10^9$ вказує на суттєвий рівень залишкової

мінливості, що є характерним для часових рядів із великими масштабами значень та значними міжрічними коливаннями.

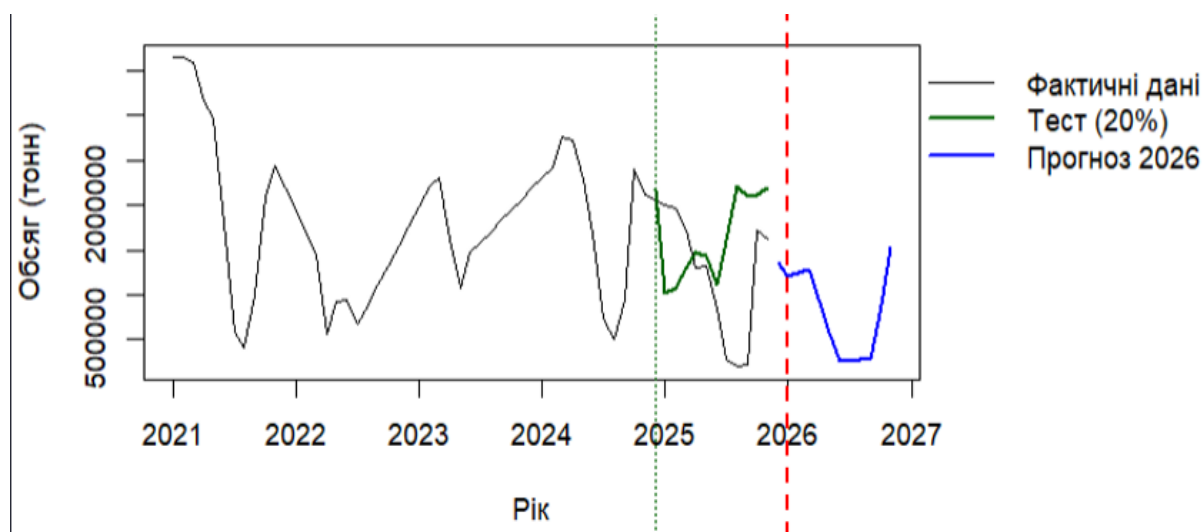


Рисунок. 2. 12 – Прогнозування обсягів реалізації кукурудзи на зерно за моделлю NNETAR

Джерело: побудовано автором

Графічне зіставлення фактичних і прогнозних значень (рис. 2.12) підтверджує зазначені особливості. Аналіз тестової вибірки, що припадає на 2025 рік, показує: хоча модель намагається повторити загальну структуру коливань і реагує на зміну тренду, вона суттєво недооцінює амплітуду падіння. Замість глибокого цільного просідання, яке спостерігається у фактичних даних, тестовий прогноз формує значно мілкішу, впадину, після чого занадто рано демонструє відскок вгору. Це свідчить про те, що неймережа схильна згладжувати екстремальні значення і не повною мірою здатна адаптуватися до різких спадів у часовому ряді. Прогноз на 2026 рік демонструє доволі нестабільну та кутасту траєкторію. Замість згладженої динаміки, NNETAR прогнозує помірне зниження на початку року, яке переходить у глибокий спад у середині періоду. Після чого очікується стрімке, майже вертикальне зростання обсягів наприкінці року.

Таблиця 2.17 – Порівняння точності прогнозних моделей для сегмента «Кукурудза на зерно» на навчальній та тестовій вибірках

Training set	ME	RMSE	MAE	MPE	MAPE	MASE
XGBOOST	828.3	13409	10 272	-0.11	0.77	0.03
ETS	11 507	359 129	287 555	-5.77	23.88	0.43
ARIMAX	-2 360	277 840	194 527	2.38	18.86	0.25
NNETAR	275.19	49 906	34 213	-0.53	3.08	0.04

Testing set	ME	RMSE	MAE	MPE	MAPE	MASE
XGBOOST	-1E+06	1E+06	980 296	-275.91	275.91	3.25
ETS	-1E+06	2E+06	1 323 305	-345.51	347.28	2.03
ARIMAX	2E+06	2E+06	1 749 442	299.23	299.23	1.62
NNETAR	-7E+05	956851	791 772	-220.73	225.82	0.73

Джерело: побудовано автором

Після детального розгляду результатів прогнозування обсягів реалізації кукурудзи на зерно за допомогою моделей ETS, XGBoost, ARIMAX та NNETAR доцільно здійснити їх узагальнене порівняння за показниками точності на навчальній і тестовій вибірках. Відповідні результати оцінювання наведено в (табл. 2.17), що дозволяє зіставити здатність моделей відтворювати історичну динаміку та їхню реальну прогнозну спроможність у позанавчальному періоді.

Аналіз результатів на навчальній вибірці показує, що найкращі значення похибок демонструє модель XGBoost, для якої зафіксовано мінімальні значення $RMSE = 13\,409$, $MAE = 10\,272$ та $MAPE = 0,77\%$. Це свідчить про високу здатність моделі точно відтворювати внутрішню структуру часового ряду кукурудзи на зерно. Модель NNETAR також демонструє відносно добрі результати ($RMSE = 49\,906$, $MAE = 34\,213$), однак поступається XGBoost за всіма ключовими показниками. Підходи ETS та ARIMAX характеризуються значно вищими значеннями похибок, що вказує на обмежену здатність цих моделей точно описувати складну та масштабну динаміку реалізації кукурудзи навіть у межах навчального періоду.

Водночас результати на тестовій вибірці демонструють різке зниження прогнозної якості для всіх розглянутих моделей. Як видно з (табл. 2.17), значення MAPE перевищують 200–300 %, що свідчить про надзвичайно високу

нестабільність прогнозів у кінцевій частині ряду. Найменші значення похибок на тестовій вибірці демонструє модель NNETAR ($RMSE = 956\,850$, $MAE = 791\,772$, $MAPE = 225,82\%$, $MASE = 0,73$), тоді як модель XGBoost, незважаючи на найкращі результати на навчальній вибірці, демонструє різке зростання похибок ($MAPE = 275,91\%$, $MASE = 3,25$). Моделі ETS та ARIMAX показують ще вищі значення похибок, що вказує на їхню низьку адаптивність до різких структурних змін у динаміці реалізації кукурудзи на зерно.

Отримані результати свідчать, що сегмент «Кукурудза на зерно» характеризується надзвичайно високою волатильністю та структурною нестабільністю, що суттєво ускладнює побудову точних короткострокових прогнозів. Жодна з розглянутих моделей не демонструє високої прогнозної точності на тестовій вибірці, що підтверджує складність цього сегмента з точки зору прогнозування.

Проведений аналіз прогнозування обсягів агровиробництва для сегментів «боби сої», «велика рогата худоба» та «кукурудза на зерно» показав, що поведінка моделей суттєво залежить від структури відповідних часових рядів та економічної специфіки сегменту.

Сегмент бобів сої характеризується вираженою сезонністю та циклічністю, що забезпечило відносно високу точність прогнозування, особливо для моделі XGBoost, яка найкраще відтворює амплітуду коливань і динаміку ряду.

Для сегмента великої рогатої худоби притаманна більш згладжена динаміка з низьким рівнем сезонності, що зумовлює ефективність моделей, орієнтованих на врахування інерційності та факторного впливу. У цьому випадку XGBoost також забезпечує найкращі результати, тоді як класичні моделі формують надто спрощені прогнози.

Сегмент кукурудзи на зерно є найбільш складним для прогнозування через високу волатильність і наявність структурних зламів, що призводить до значного зниження точності всіх моделей на тестовій вибірці.

Загалом результати дослідження підтверджують, що універсальної моделі прогнозування для всіх сегментів агровиробництва не існує. Найбільш ефективним

є використання багатомодельного підходу з урахуванням специфіки динаміки кожного сегменту, при якому XGBoost доцільно застосовувати для точного відтворення складних залежностей, тоді як ETS та ARIMAX можуть виконувати роль базових і пояснювальних інструментів.

2.3. Кластеризація видів сільськогосподарської продукції

Наступним етапом дослідження є кластеризація видів сільськогосподарської продукції. Кластерний аналіз дає змогу перейти до виявлення внутрішньої структури всієї сукупності досліджуваних об'єктів. Його застосування дозволяє згрупувати види продукції за ступенем подібності їх виробничих, ресурсних і ринкових характеристик.

Необхідність застосування кластеризації також впливає з результатів підрозділу 2.1, де було встановлено значну неоднорідність досліджуваного масиву даних. Зокрема, показники обсягів реалізації, площ посівів, рівня цін, використання добрив та інших характеристик істотно різняться за масштабом, мають асиметричні розподіли та містять викиди. Крім того, кореляційний аналіз показав наявність окремих змістовних зв'язків між ознаками за відсутності критичної мультиколінеарності, що створює належні передумови для застосування багатовимірних методів групування. Саме тому кластеризація в межах даної роботи розглядається як інструмент систематизації різномірних видів аграрної продукції та виявлення латентної структури агровиробництва України.

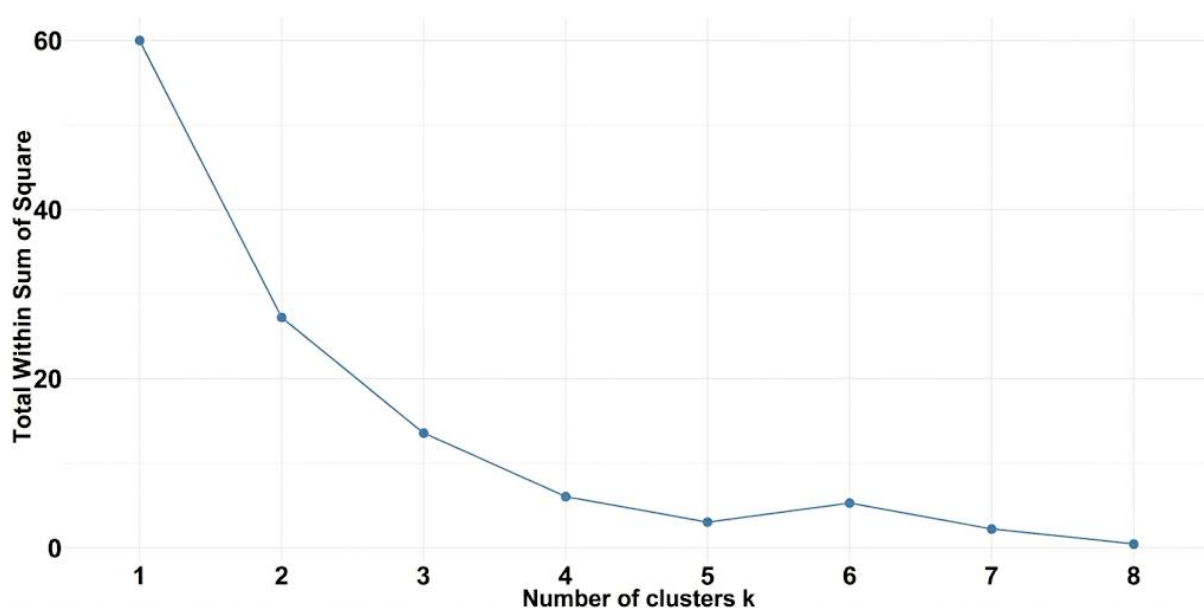


Рисунок 2.18 – Обґрунтування вибору кількості кластерів за методом ліктя

Джерело: побудовано автором

Першим етапом кластерного аналізу стало визначення оптимальної кількості кластерів (рис. 2.18). Для цього було використано метод ліктя, який базується на аналізі зміни внутрішньокластерної суми квадратів відстаней при послідовному збільшенні кількості кластерів. За результатами побудови відповідного графіка встановлено, що найбільш обґрунтованим є вибір чотирьох кластерів, оскільки саме при $k=4$ спостерігається характерне уповільнення зменшення внутрішньокластерної дисперсії. Це означає, що подальше збільшення кількості кластерів не забезпечує істотного приросту якості групування.

Таблиця 2.18 – Характеристика кластерів основних видів сільськогосподарської продукції за результатами кластеризації k-means

№ Кластеру	Кількість культур	Склад кластера	Сер. Обсяг	Сер. Ціна	Сер. Площа (га)	Сер. Обсяг Добрив	Сер. Поголов'я
1	5	Боби сої, Жито, Картопля, Культури овочеві, Ячмінь	232 585.3	7 764.3	1 445 522	520.6	0
2	3	Кукурудза на зерно, Культури зернові та зернобобові, Пшениця	2 621 034	5 005.5	23 710 628	16 077.5	0
3	1	Птиця свійська	75 973.3	44 156.6	0	0	1.02E+09
4	3	Велика рогата худоба, Свині, Тварини сільськогосподарські живі	60 738.3	38 398.6	0	0	71 929 669

Джерело: побудовано автором

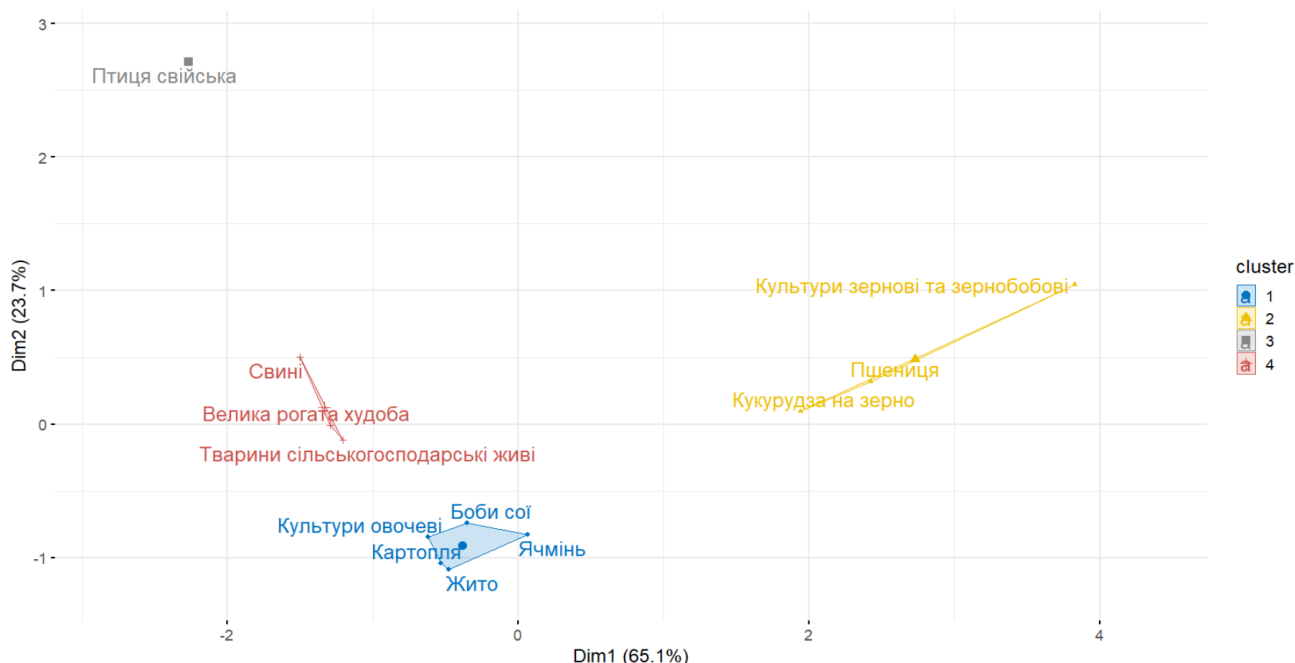


Рисунок 2.19 – Результати кластеризації основних видів сільськогосподарської продукції методом k-means

Джерело: побудовано автором

У результаті застосування алгоритму *k-means* до 12 основних видів сільськогосподарської продукції було виділено чотири відносно однорідні кластери, кожен з яких характеризується власним поєднанням виробничих, ресурсних і ринкових параметрів (рис.2.19)(табл.2.18). До першого кластера увійшли боби сої, жито, картопля, культури овочеві та ячмінь. Для цього сегмента характерні середній обсяг реалізації на рівні 232 585,3 т, середня ціна 7 764,3 грн/т, середня площа 1 445 522 га та середній обсяг добрив 520,6 т. Відсутність показника поголів'я є закономірною, оскільки до цього кластера входять виключно рослинницькі види продукції. За своїм економічним змістом цей кластер доцільно інтерпретувати як сегмент рослинницьких культур переважно внутрішнього ринку та змішаного споживчо-сировинного призначення, що характеризується помірними масштабами виробництва та відносно нижчим рівнем ресурсної інтенсивності порівняно з найбільшими зерновими культурами.

Другий кластер об'єднав кукурудзу на зерно, культури зернові та зернобобові, а також пшеницю. Саме цей кластер виявився найбільш масштабним за ключовими параметрами: середній обсяг реалізації становить 2 621 034 т,

середня ціна — 5 005,5 грн/т, середня площа — 23 710 628 га, а середній обсяг використання добрив — 16 077,5 т. Такі характеристики свідчать про домінування в межах цього кластера великотоварного, масового та експортно орієнтованого зернового виробництва. Порівняно з першим кластером тут спостерігається кратно більший земельний банк, вищий рівень ресурсного забезпечення та значно більші фізичні обсяги реалізації, що відображає стратегічну роль зернової групи у формуванні аграрного виробництва та експортного потенціалу України.

Третій кластер представлений лише одним об'єктом — продукцією категорії «птиця свійська». Незважаючи на формальну одноелементність, виділення цього окремого кластера є цілком закономірним, оскільки птахівництво суттєво відрізняється від інших видів продукції за поєднанням економічних характеристик. Для нього характерні середній обсяг реалізації 75 973,3 т, середня ціна 44 156,6 грн/т та надзвичайно велике середнє поголів'я — 1 023 285 858 голів. Відсутність площ посівів і добрив у структурі цього кластера також є логічною, оскільки відповідні показники не є базовими для оцінювання тваринницького напрямку. З економічної точки зору цей кластер доцільно трактувати як окремий сегмент індустріального птахівництва, що характеризується високою концентрацією виробництва, інтенсивним оборотом капіталу та відносно вищою ціною реалізації порівняно з класичними сегментами тваринництва.

Четвертий кластер сформували велика рогата худоба, свині та тварини сільськогосподарські живі. Для нього характерні середній обсяг реалізації 60 738,3 т, середня ціна 38 398,6 грн/т та середнє поголів'я 71 929 669 голів. Як і у випадку третього кластера, показники площ та добрив тут дорівнюють нулю, що відображає специфіку тваринницького напрямку та відмінність від рослинництва. За змістом цей кластер можна визначити як сегмент традиційного тваринництва, де поєднуються довший виробничий цикл, нижча оборотність капіталу порівняно з птахівництвом і водночас достатньо високий рівень цін реалізації. У порівнянні з кластером індустріального птахівництва цей сегмент є менш масштабним за поголів'ям та обсягами реалізації, проте зберігає вагоме місце у структурі агровиробництва як окремий тип господарської спеціалізації.

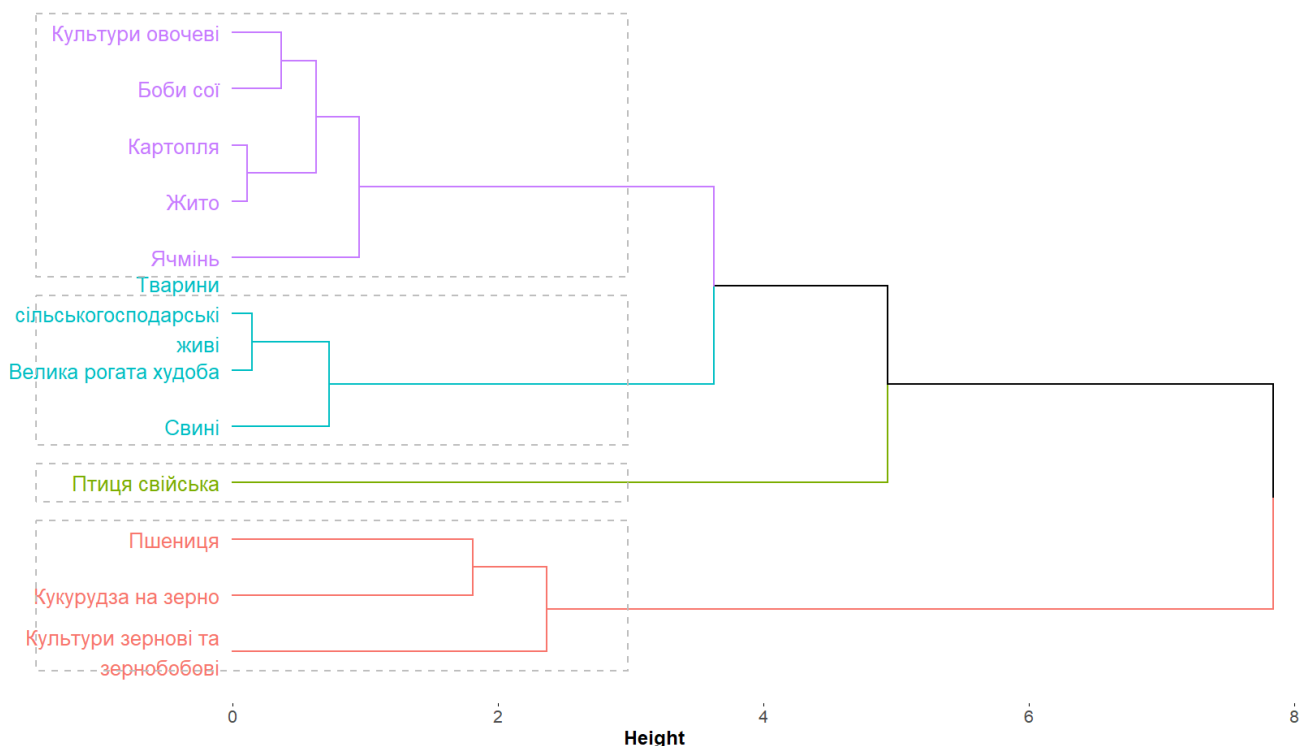


Рисунок 2.20 – Дендрограма ієрархічної кластеризації основних видів сільськогосподарської продукції

Джерело: побудовано автором

Для перевірки стійкості результатів кластеризації додатково було застосовано метод ієрархічної кластеризації. Отримана дендрограма (рис. 2.20) підтверджує наявність чотирьох основних груп, які відповідають результатам методу k-means. Зокрема, чітко простежується відокремлення зернових культур, тваринницького сегмента та групи продуктів внутрішнього ринку, що свідчить про узгодженість отриманих результатів різними методами.

Виділення окремого кластеру експортно орієнтованих зернових культур свідчить про високу концентрацію земельних і ресурсних потоків у небагатьох сегментах агровиробництва. Зокрема, другий кластер демонструє найбільші середні площі посівів та найвищі обсяги використання добрив, що вказує на підвищену ресурсну інтенсивність такого типу виробництва. Натомість перший кластер характеризується значно меншими масштабами площ і добрив, що може свідчити про іншу модель функціонування — менш капіталоємну, але й менш масштабну з точки зору фізичних обсягів реалізації. У тваринницькому секторі, своєю чергою, кластеризація показує, що навіть за відсутності земельної складової

існують суттєві відмінності між різними видами тваринницької продукції за рівнем індустріалізації та ціновими характеристиками. Це дозволяє розглядати стійкість агровиробництва не як єдину універсальну категорію, а як множину різних сегментів, кожен з яких має власні обмеження, ризики та потенціал розвитку.

Застосування кластерного аналізу дозволило виділити чотири типологічно відмінні сегменти сільськогосподарської продукції: рослинницькі культури помірного масштабу, великотоварне зернове виробництво, індустріальне птахівництво та традиційне тваринництво. Отримане групування є аналітично змістовним, економічно інтерпретованим і таким, що узгоджується як з попередньо обґрунтованою методикою кластеризації, так і з результатами статистичного та прогнозного аналізу. Практична цінність цього результату полягає в тому, що він створює основу для подальшої диференційованої оцінки сталого розвитку агровиробництва: не на рівні абстрактної середньої галузі, а на рівні конкретних структурних типів, які мають різну ресурсну інтенсивність, різні ринкові параметри та різні траєкторії функціонування.

ВИСНОВКИ

У кваліфікаційній бакалаврській роботі досягнуто поставленої мети – розроблено та практично апробовано математичні підходи до оцінювання сталого розвитку агровиробництва України на основі статистичного аналізу та сучасних методів моделювання.

У процесі дослідження уточнено сутність агровиробництва та розглянуто теоретичні засади його сталого розвитку. Показано, що агровиробництво доцільно розглядати як складну багаторівневу соціально-економічну систему, в якій поєднуються виробничі, ринкові, ресурсні, екологічні та інституційні складові. Встановлено, що оцінювання його сталого розвитку має здійснюватися з урахуванням взаємодії економічної результативності, екологічної безпеки та соціальної складової. Окремо враховано сучасні українські умови, зокрема вплив воєнних ризиків, скорочення площ посівів, погіршення ресурсної бази та дефіцит трудових ресурсів.

Під час аналізу існуючих підходів до оцінювання сталого розвитку аграрного сектору було з'ясовано, що суто теоретичні моделі вже не є достатніми. Більш обґрунтованим виглядає поєднання різних підходів – статистичного аналізу, економіко-математичного моделювання та методів машинного навчання. Такий комплексний підхід дозволяє краще враховувати як часову мінливість показників, так і вплив зовнішніх факторів та структурну неоднорідність галузі.

Відповідно до поставлених завдань обґрунтовано вибір інструментарію дослідження. Для прогнозування використано моделі ETS, ARIMAX, NNETAR та XGBoost, що дозволило порівняти класичні та сучасні підходи до аналізу часових рядів. Для аналізу структурної неоднорідності застосовано метод ліктя, алгоритм k-means та ієрархічну кластеризацію. Така комбінація методів забезпечує можливість комплексного аналізу з використанням системи показників точності (ME, RMSE, MAE, MAPE, MASE).

У межах роботи сформовано статистичний датасет на основі офіційних джерел. До нього включено показники, що характеризують різні аспекти агровиробництва: обсяги реалізації, ціни, площі, добрива, інвестиції та макроекономічні параметри. Під час попереднього аналізу стало очевидно, що показники мають різні масштаби та значну варіативність, а також різну повноту спостережень. Це зумовило необхідність їх попередньої обробки та стандартизації.

Результати статистичного аналізу підтвердили, що досліджувані дані є неоднорідними. Для більшості змінних характерні асиметрія розподілів та наявність викидів. Кореляційний аналіз не виявив критичної мультиколінеарності, хоча окремі зв'язки є економічно змістовними. Додатково встановлено, що для багатьох видів продукції фактор часу є статистично значущим, а структура часових рядів суттєво відрізняється за співвідношенням трендової, сезонної та випадкової компонент.

У частині прогнозування проведено порівняння кількох моделей для різних сегментів агровиробництва. Отримані результати показали, що універсального підходу, який однаково добре працює для всіх видів продукції, не існує. Водночас загалом найбільш ефективною виявилась модель XGBoost, яка забезпечила кращу точність на тестових даних. Зокрема, для сегмента «Боби сої» та для сегмента «Велика рогата худоба». Водночас для кукурудзи на зерно всі моделі продемонстрували низьку точність, що пов'язано з високою волатильністю цього сегмента.

Окрему увагу приділено кластерному аналізу. У результаті застосування методу k-means та ієрархічної кластеризації було виділено чотири групи продукції, які характеризують різні типи агровиробництва: культури внутрішнього ринку, великотоварне зернове виробництво, індустриальне птахівництво та традиційне тваринництво. Це підтверджує, що аграрний сектор України є структурно неоднорідним, а отже, його доцільно аналізувати не лише в цілому, а й з урахуванням окремих типів спеціалізації.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Огляд ринку зернових культур в Україні: поточний стан і перспективи галузі в контексті глобальних трендів 2025 : аналітичний звіт. Сайт BDO в Україні. 2025. URL: <https://www.bdo.ua/uk-ua/insights-2/information-materials/2025/rynok-zernovukh-ukrayina-2025> (дата звернення: 20.03.2026).
2. Ажнюк М. О., Передрій О. С. Аграрні відносини та їх система. Місце аграрних відносин у ринковій економіці. Основи економічної теорії : навч. посіб.
3. Богонос М., Чміль А., Мартишев П. та ін. Агровигляд України 2024-2033 : матеріали доповіді. Сайт Центру досліджень продовольства та землекористування KSE. 2024. URL: https://kse.ua/wp-content/uploads/2024/03/Agroviglyad_2_ukr.pdf (дата звернення: 21.03.2026).
4. Економіка та підприємництво : зб. наук. праць. № 38 / голов. ред. І. М. Рєпіна. Київ : КНЕУ, 2017. 334 с.
5. Державна служба статистики України. Бізнес-статистика: Сільське, лісове та рибне господарство. Офіційний сайт Державної служби статистики України. URL: <https://stat.gov.ua/uk/explorer> (дата звернення: 20.03.2026).
6. Макроекономічні показники: Сільське господарство. Київ : Офіц. сайт НБУ. URL: <https://bank.gov.ua/ua/statistic/macro-indicators#2> (дата звернення: 20.03.2026).
7. Русан В. М., Жураковська Л. А. Аграрний сектор України у 2023 році: складники стійкості, проблеми та перспективні завдання. Сайт Національного інституту стратегічних досліджень, 2024. URL: https://niss.gov.ua/sites/default/files/2024-02/az_agrosektor_15022024.pdf (дата звернення: 21.03.2026).
8. Вплив сезонності та біологічних циклів на управління виробничими процесами сільськогосподарських підприємств / В. Ніценко, Н. Присяжнюк, Ю. Федорук та ін. Економічний аналіз. 2025. Т. 35, № 4. С. 259–267. URL:

- <https://www.econa.org.ua/index.php/econa/article/view/6504/6565657779> (дата звернення: 21.03.2026).
9. Олійник Т. Г., Мельник І. О., Горобченко О. А. Економіка аграрного виробництва : курс лекцій. Миколаїв : МНАУ, 2019. 105 с.
 10. Русан В. М. Особливості функціонування аграрного сектора економіки України в умовах війни. Сайт Національного інституту стратегічних досліджень, 2022. URL: https://niss.gov.ua/sites/default/files/2024-02/az_agrosector_15022024.pdf (дата звернення: 21.03.2026).
 11. Садовська І. Б., Замула І. В., Мельник К. П. Політика формування забезпечень та резервів в обліково-аналітичній системі сільськогосподарського підприємства. Економіка та суспільство. 2026. № 83. URL: <https://economyandsociety.in.ua/index.php/journal/article/view/7518/7483> (дата звернення: 21.03.2026).
 12. Сікачина О., Гагалюк Т. Аналітичний огляд структури аграрного сектора України в контексті європейської інтеграції : агрополітичний звіт APD/APB/05/2025. URL: https://www.apd-ukraine.de/fileadmin/user_upload/APD_Bericht_UA_251029_OS_FINAL.pdf (дата звернення: 21.03.2026).
 13. Трегобчук В. М. Концепція сталого розвитку для України. Вісник Національної академії наук України. 2002. № 2. С. 31–40. URL: http://nbuv.gov.ua/UJRN/vnanu_2002_2_7 (дата звернення: 21.03.2026).
 14. Cabezas L.M.C., Izbicki R., Stern R.B. Hierarchical clustering: visualization, feature importance and model selection. arXiv, 2023. URL: <https://arxiv.org/pdf/2112.01372> (date of access: 21.03.2026).
 15. Cobourn K. Climate change adaptation policies to foster resilience in agriculture: Analysis and stocktake based on UNFCCC reporting documents : OECD Food, Agriculture and Fisheries Papers No. 202. OECD Publishing, 2023. URL: <https://reliefweb.int/report/world/climate-change-adaptation-policies-foster-resilience-agriculture-oecd-food-agriculture-and-fisheries-paper-july-2023-ndeg202> (date of access: 21.03.2026).

16. Albaladejo Román A. Ukrainian agriculture: From Russian invasion to EU integration : briefing PE 760.432. European Parliamentary Research Service, 2024. URL: https://www.europarl.europa.eu/RegData/etudes/BRIE/2024/760432/EPRS_BRI%282024%29760432_EN.pdf (date of access: 21.03.2026).
17. Federated XGBoost. Federated XGBoost model: Official documentation. Read the Docs. URL: <https://federated-xgboost.readthedocs.io/en/latest/tutorials/model.html> (date of access: 21.03.2026).
18. Food and Agriculture Organization. Increasing crop production sustainably: The perspective of biological processes. Rome : FAO, 2009. URL: <https://openknowledge.fao.org/server/api/core/bitstreams/27e0746a-1577-4643-bcf9-57c8506dbe73/content> (date of access: 22.03.2026).
19. Ukraine: FAO surveys impact of war on agricultural enterprises. Official website of FAO, 2025. URL: <https://www.fao.org/countryprofiles/news-archive/detail-news/en/c/1742727/> (date of access: 22.03.2026).
20. The State of Food and Agriculture 2023 – Revealing the true cost of food to transform agrifood systems. Rome : FAO, 2023. URL: <https://openknowledge.fao.org/server/api/core/bitstreams/5aac5078-625d-4b94-b964-bea40493016c/content> (date of access: 22.03.2026).
21. Environmental Sustainability in Agriculture 2023. Rome : OECD/FAO, 2023. URL: <https://openknowledge.fao.org/server/api/core/bitstreams/32da2942-3854-4736-af19-877b3ab22d35/content> (date of access: 22.03.2026).
22. What is an ARIMAX model? GeeksforGeeks. URL: <https://www.geeksforgeeks.org/artificial-intelligence/what-is-an-arimax-model/> (date of access: 22.03.2026).
23. HLPE. Executive summary of the report Building resilient food systems. Rome : FAO, 2025. URL: https://sfcs.fao.org/media-cnt/docs/devhlpelibraries/report-20/hlpe20_executive-summary_en_web.pdf (date of access: 22.03.2026).
24. Hyndman R. J., Athanasopoulos G. Forecasting with ETS models. Forecasting: Principles and practice. 2nd ed. OTexts, 2018. URL: <https://otexts.com/fpp2/ets.html> (date of access: 22.03.2026).

25. Hyndman R. J., Athanasopoulos G. Neural network autoregression. Forecasting: Principles and practice. 3rd ed. OTexts, 2021. URL: <https://otexts.com/fpp3/nnetar.html> (date of access: 22.03.2026).
26. Kmytiuk T., Majore G., Bilyk T. Time series forecasting of price of the agricultural products using data science. Agricultural and Resource Economics: International Scientific E-Journal, 2024. Vol 10(3) P. 5-33. URL: <https://are-journal.com/are/article/view/906> (date of access: 22.03.2026).
27. Neyter R., Zorya S., Muliar O. Agricultural War Damages, Losses, and Needs Review. Kyiv : KSE Center for Food and Land Use Research, 2024. URL: https://fscluster.org/sites/default/files/2024-05/RDNA3_eng.pdf (date of access: 22.03.2026).
28. OECD, FAO. Environmental Sustainability in Agriculture 2023 : issues note. Rome : OECD/FAO, 2023. URL: <https://www.oecd.org/content/dam/oecd/en/topics/policy-issues/agriculture-and-sustainability/environmental-sustainability-in-agriculture-OECD-FAO-issues-note-2023.pdf> (date of access: 22.03.2026).
29. Pyvovar P., Rozhkov O. Current state and trends of cattle rearing, cattle meat and milk sectors in Ukraine : Eurointegration 1.0: Sector overview. KSE Center for Food and Land Use Research, 2025. URL: https://kse.ua/wp-content/uploads/2025/02/Cattle_UA.pdf (date of access: 22.03.2026).
30. Solokha M., Demyanyuk O., Symochko L. et al. Soil Degradation and Contamination Due to Armed Conflict in Ukraine. Land, 2024. Vol. 13, No. 10. 1614. URL: <https://www.mdpi.com/2073-445X/13/10/1614> (date of access: 22.03.2026).
31. Tarashevych O. Livestock and Products Annual: Ukraine : GAIN Report Number: UP2024-0017. USDA Foreign Agricultural Service, 2024. URL: https://apps.fas.usda.gov/newgainapi/api/Report/DownloadReportByFileName?fileName=Livestock%20and%20Products%20Annual_Kyiv_Ukraine_UP2024-0017.pdf (date of access: 22.03.2026).
32. AutoElbow: An Automatic Elbow Detection Method for Estimating the Number of Clusters in a Dataset / A. J. Onumanyi, D. N. Molokomme, S. J. Isaac, A. M. Abu-

Mahfouz. *Applied Sciences*, 2022. Vol. 12, No. 15. 7515. URL: <https://www.mdpi.com/2076-3417/12/15/7515> (date of access: 22.03.2026).

33. Seeds of hope: helping Ukraine farmers build resilience. 03.10.2024. Official website of the World Bank. URL: <https://www.worldbank.org/en/news/immersive-story/2024/10/03/seeds-of-hope-helping-ukraine-farmers-build-resilience> (date of access: 22.03.2026).

Метадані

ДОКУМЕНТ

Заголовок

Власов Кмитюк плагіат

Автор

Власов

Науковий керівник / Експерт

Кмитюк

ІД документу

334099026

ОРГАНІЗАЦІЯ

Назва організації

**Kyiv National Economic University named after Vadym Hetman
KNEU**

підрозділ

кафедра штучного інтелекту, моделювання та статистики

ЗВІТ

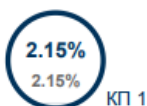
Дата звіту

5/30/2026

Дата редагування

Обсяг знайдених подібностей

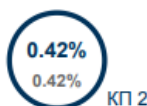
Коефіцієнт подібності визначає, який відсоток тексту по відношенню до загального обсягу тексту було знайдено в різних джерелах. Зверніть увагу, що високі значення коефіцієнта не автоматично означають плагіат. Звіт має аналізувати компетентна / уповноважена особа.



КП 1

25

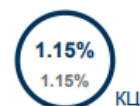
Довжина фрази для коефіцієнта подібності 2



КП 2

12641

Кількість слів



КЦ

98967

Кількість символів

Тривога

У цьому розділі ви знайдете інформацію щодо текстових спотворень. Ці спотворення в тексті можуть говорити про МОЖЛИВІ маніпуляції в тексті. Спотворення в тексті можуть мати навмисний характер, але частіше характер технічних помилок при конвертації документа та його збереженні, тому ми рекомендуємо вам підходити до аналізу цього модуля відповідально. У разі виникнення запитань, просимо звертатися до нашої служби підтримки.

Заміна букв		16
Інтервали		0
Мікропробіли		106
Білі знаки		0
Парафрази (SmartMarks)		12

Джерела

Нижче наведений список джерел. В цьому списку є джерела із різних баз даних. Колір тексту означає в якому джерелі він був знайдений. Ці джерела і значення Коефіцієнту Подібності не відображають прямого плагіату. Необхідно відкрити кожне джерело і проаналізувати зміст і правильність оформлення джерела.

10 найдовших фраз

Копіювати текст

#	НАЗВА ТА АДРЕСА ДЖЕРЕЛА URL (НАЗВА БАЗИ)	КІЛЬКІСТЬ ІДЕНТИЧНИХ СЛІВ (ФРАГМЕНТІВ)
1	https://www.tandfonline.com/doi/full/10.1080/08839514.2022.2074129	28 (0 22 %)

Кмитюк Т.Л.

к.е.н., доцент,

Власов Д.А.

здобувач першого (бакалаврського) рівня вищої освіти,

Київський національний економічний університет імені Вадима Гетьмана

МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ СТАЛОГО РОЗВИТКУ АГРОВИРОБНИЦТВА: ПОРІВНЯЛЬНИЙ АНАЛІЗ ПРОГНОЗНИХ МОДЕЛЕЙ ДЛЯ ОКРЕМИХ СЕГМЕНТІВ

Агровиробництво є однією з ключових галузей економіки України, оскільки забезпечує продовольчу безпеку, формує значну частку експортного потенціалу та суттєво впливає на макроекономічну стабільність. Водночас сучасний розвиток аграрного сектору відбувається в умовах високої невизначеності, зумовленої воєнними ризиками, ресурсними обмеженнями, екологічними викликами та нестабільністю ринкової кон'юнктури. Це посилює потребу у використанні стандартизованих підходів до оцінювання та прогнозування динаміки агровиробництва.

Метою дослідження є порівняльне оцінювання прогнозних моделей для окремих сегментів агровиробництва з позицій їх придатності до аналізу сталого розвитку. У роботі використано сформований датасет на основі офіційних статистичних джерел [4].

Для дослідження були застосовані чотири групи моделей: ETS як класичний інструмент експоненційного згладжування, ARIMAX як економетричний підхід із зовнішніми змінними, XGBoost як алгоритм машинного навчання для врахування складних нелінійних залежностей і NNETAR як нейромережева модель часових рядів [2,3]. Для оцінювання точності прогнозів використовувалися показники RMSE, MAE, MAPE та MASE, а часові ряди поділялися на навчальну і тестову вибірки у співвідношенні 80/20 [1]. Такий підхід дозволив не лише оцінити здатність моделей відтворювати історичну динаміку, а й перевірити їх стійкість у позанавчальному періоді.

Дослідження було зосереджено на трьох сегментах: жито, картопля та культури зернові. Саме ці сегменти добре демонструють різну часову структуру рядів і, відповідно, різну поведінку моделей.

Таблиця 1

Оцінювання точності моделі на навчальних і тестових даних сегмента «Жито»

Training set	ME	RMSE	MAE	MPE	MAPE	MASE
XGBOOST	2.4052	166.58	129.63	0.3782	1.555388	0.020543
ETS	-153.5	2017.4	1516.8	-6.704	23.81458	0.368992
ARIMAX	-458.8	2961	2010.6	-8.331	28.08349	0.236911
NNETAR	-7.68	788.66	583.19	-3.841	9.801031	0.068719
Testing set	ME	RMSE	MAE	MPE	MAPE	MASE
XGBOOST	2600.1	5305.6	4340.9	24.508	85.75594	0.687939
ETS	-39.31	4811.1	3569.3	-56.43	80.08	0.754
ARIMAX	-1430	4864.5	3945.5	-87.27	106.66	0.368
NNETAR	-8484	9951.3	9168.3	-279.2	282.98	0.855

Результати прогнозування сегмента «Жито» показали, що жодна з моделей не забезпечує ідеальної точності на всіх етапах аналізу. На навчальній вибірці найкращі показники продемонструвала модель XGBoost, для якої RMSE становив 166,58, MAE — 129,63, а MAPE — 1,56 %, що свідчить про високу здатність моделі відтворювати внутрішню структуру ряду. На тестовій вибірці більш стабільними виявилися ETS та ARIMAX: для ETS MAPE становив 80,08 %, а для ARIMAX — 106,66 %, тоді як NNETAR показала найгіршу якість узагальнення з MAPE понад 280 %. Для сегмента жита моделі ETS та ARIMAX є більш прийнятними з позицій практичного прогнозування, тоді як XGBoost, попри високу точність на етапі навчання, виявляє ознаки переналаштування на історичні дані.

Таблиця 2

Оцінювання точності моделі на навчальних і тестових даних сегмента «Картопля»

Training set	ME	RMSE	MAE	MPE	MAPE	MASE
XGBOOST	7.68408	104.3	85.6562	-0.05286	0.5286843	0.02402
ETS	73.3086	4719	3512.08	-3.51571	20.79613	0.69637
ARIMAX	-13.295	5310	3758.86	-6.74134	19.4538	0.42258
NNETAR	-5.0905	1509	1018.19	-1.51271	6.18009	0.11447

Продовження таблиці 2

Testing set	ME	RMSE	MAE	MPE	MAPE	MASE
XGBOOST	3278.86	11636	9315.63	-7.19912	45.129675	2.61279
ETS	-2353.6	13453	11852.1	-44.16	65.19	3.568
ARIMAX	2877.72	12391	9135.5	-11.65	42.17	0.883
NNETAR	-2900.1	11731	9564.4	-39.07	56.13	0.924

Аналіз сегмента картоплі виявив дещо інші тенденції. Найкращу точність на навчальній вибірці знову забезпечила модель XGBoost: RMSE = 104,3, MAE = 85,66, MAPE = 0,53 %. Проте на тестовій вибірці найбільш збалансованою виявилася модель ARIMAX, яка забезпечила MAPE = 42,17 % та MASE = 0,88. Модель XGBoost у цьому сегменті хоч і зберегла відносно добрі результати, однак поступилася ARIMAX за стабільністю прогнозу поза межами періоду навчання. Для картоплі можна зробити висновок про доцільність поєднання класичних економетричних підходів із машинним навчанням: XGBoost добре відтворює історичну динаміку, а ARIMAX краще узагальнює її у короткостроковому прогнозі.

Таблиця 3

Оцінювання точності моделі на навчальних і тестових даних сегмента «Зернові та зернобобові культури»

Training set	ME	RMSE	MAE	MPE	MAPE	MASE
XGBOOST	932.3579	9991	7615.6	-0.0191	0.3045	0.020991
ETS	12768.63	6E+05	465621	-2.6449	18.593	0.436019
ARIMAX	-22662.59	5E+05	345312	-3.4559	14.9	0.323359
NNETAR	-39.29688	11553	6201	-0.028	0.2678	0.005807
Testing set	ME	RMSE	MAE	MPE	MAPE	MASE
XGBOOST	-756602.43	1E+06	877136	-42.23	45.544	2.41772
ETS	231427.78	6E+05	502689	9.69	20.05	0.392
ARIMAX	2205138.14	2E+06	2E+06	87.1	87.1	1.721
NNETAR	-1360808.0	1E+06	1E+06	-64.88	64.88	1.062

Специфіка прогнозування для агрегованої групи зернових та зернобобових культур суттєво відрізняється. На навчальній вибірці найкращі результати продемонстрували NNETAR і XGBoost: MAPE для NNETAR становив 0,27 %, а для XGBoost — 0,30 %, що вказує на високу здатність обох моделей адаптуватися до внутрішньої структури агрегованого ряду. Однак на тестовій вибірці найменші похибки забезпечила модель ETS, для якої MAPE становив 20,05 %, тоді як

XGBoost мав MAPE 45,54 %, NNETAR — 64,88 %, а ARIMAX — 87,10 %. Це свідчить про те, що для агрегованих показників, де індивідуальні коливання частково згладжуються, класичні статистичні методи можуть виявитися більш надійними, ніж складні машинно-навчальні або нейромережеві підходи.

Отримані результати підтверджують, що універсальної моделі прогнозування для всіх сегментів агровиробництва не існує. Вибір інструментарію повинен залежати від характеру часового ряду, ступеня сезонності, впливу зовнішніх факторів та мети аналізу. Для сегментів із вираженою нелінійністю та складною внутрішньою структурою XGBoost часто демонструє найкращі результати на етапі навчання. Для більш стабільних або агрегованих рядів ETS і ARIMAX можуть забезпечувати вищу надійність прогнозу в практичному застосуванні. Нейромережеві моделі NNETAR доцільно розглядати як допоміжний інструмент, особливо для виявлення нелінійних ефектів, але їх використання потребує обережної інтерпретації.

З позицій оцінювання сталого розвитку агровиробництва запропонований підхід має практичну цінність, оскільки дозволяє враховувати не лише поточні обсяги реалізації, а й їхню чутливість до ресурсного забезпечення, цінових змін, інвестиційної активності та макроекономічного середовища. Це створює основу для більш обґрунтованого прийняття управлінських рішень як на рівні окремих виробників, так і на рівні державної аграрної політики. Поєднання статистичних, економетричних і машинно-навчальних методів дає змогу підвищити якість аналітичної оцінки аграрного сектору України в умовах невизначеності та структурних зрушень.

Проведене порівняння доводить, що відмова від універсальних прогнозних моделей на користь сегментованого підходу (XGBoost для складних нелінійних даних, ARIMAX/ETS для агрегованих) є необхідною умовою для коректного аналізу сталого розвитку агровиробництва. Саме багатомодельний підхід забезпечує найбільш коректну інтерпретацію процесів, що відбуваються в агровиробництві України, та формує основу для подальших досліджень у сфері цифрової економіки, аналітики даних і підтримки управлінських рішень.

Список використаних джерел

1. Kmytiuk, T., Majore, G., & Bilyk, T. (2024). Time series forecasting of price of the agricultural products using data science. *Agricultural and Resource Economics*, 10(3), 5–33. <https://doi.org/10.51599/are.2024.10.03.01>
2. Rob J Hyndman, & George Athanasopoulos. (2021). *Forecasting: Principles and practice* (3rd ed.). OTexts.
3. Torres, R., Ohashi, O., Pessanha, S., et al. (2022). A review of time series forecasting models. *Applied Sciences*, 12(15), 7515.
4. Державна служба статистики України. (n.d.). *Сільське господарство, лісове господарство та рибне господарство*. <https://stat.gov.ua/>