

А. Р. Ліщук

ПОРІВНЯЛЬНИЙ АНАЛІЗ СТАТИСТИЧНИХ ПІДХОДІВ І НЕЙРО-НЕЧІТКИХ МОДЕЛЕЙ ДЛЯ ПРОГНОЗУВАННЯ ТЕМПЕРАТУРИ ЗЕРНА

Забезпечення оптимальних умов зберігання зерна є важливою задачею аграрного сектору, оскільки відхилення температурного режиму може спричинити значні економічні втрати через псування зернової маси. У статті досліджено ефективність статистичних та нейро-нечітких методів прогнозування температури зерна у зерносховищах. Розглянуто основні статистичні підходи, зокрема лінійну та поліноміальну регресію, методи ARIMA та експоненційного згладжування, які широко використовуються для аналізу часових рядів. Також проаналізовано можливості нейро-нечітких моделей, таких як ANFIS, що базується на нечіткій моделі Sugeno, та Mamdani. Ці моделі поєднують переваги штучних нейронних мереж та нечіткої логіки, дозволяючи адаптивно враховувати складні нелінійні взаємозв'язки між параметрами.

Для порівняння точності прогнозування різними підходами використано реальні та симуляційні дані температурних змін у зерносховищах. Прогнозування здійснювалося для температури в кожній із 200 вимірювальних точок. Використано методику ковзного вікна (14 днів) для побудови часових рядів. Оцінка ефективності моделей здійснювалася за допомогою середньоквадратичної похибки (MSE), середньої абсолютної похибки (MAE) та кореня середньоквадратичної похибки (RMSE), з урахуванням діапазону змін температури у сховищі.

Результати експериментального дослідження показали, що статистичні методи, особливо ARIMA та експоненційне згладжування, є ефективними для короткострокового прогнозування (до 7 днів), тоді як їхня точність суттєво знижується на довших горизонтах. Нейро-нечіткі моделі, особливо ANFIS, продемонстрували високу точність прогнозування навіть у довгостроковій перспективі, що підтверджує їхню перевагу в умовах складної та динамічної зміни факторів, що впливають на температуру зерна.

Крім того, виконано аналіз обчислювальної складності методів. Встановлено, що статистичні підходи є менш ресурсозатратними, що робить їх придатними для вбудованих систем моніторингу, тоді як нейро-нечіткі моделі потребують значних обчислювальних потужностей, але забезпечують вищу точність прогнозів. На основі отриманих результатів сформульовано рекомендації щодо вибору оптимального методу прогнозування залежно від специфіки застосування, доступних обчислювальних ресурсів та вимог до точності прогнозів.

Ключові слова: прогнозування температури зерна, регресійні моделі, нейро-нечіткі моделі, порівняльний аналіз, точність прогнозування.

Постановка задачі

Забезпечення належних умов зберігання зерна є одним із ключових завдань аграрної галузі, оскільки зміни температури в зерносховищах можуть спричинити псування зернової маси, розвиток шкідників і мікроорганізмів, що призводить до значних економічних втрат. За даними досліджень, підвищення температури зерна понад критичний рівень у сховищах без належного контролю може призвести до його самозігрівання, що створює ризик займання [1]. Крім того, нерівномірний розподіл температури всередині зернової маси сприяє локальному утворенню зон підвищеної вологості, що значно прискорює розвиток грибкових захворювань та зниження якості продукції [2].

У цьому дослідженні передбачено прогнозування температури зерна в умовах його зберігання у зерносховищі. Відомими є параметри, що реєструються датчиками: температура навколишнього середовища (°C), вологість повітря (%), вологість зерна (%), щільність насипу (кг/м³), рівень вентиляції (м³/год), початкова температура зерна (°C), час зберігання (добы), тип зернової культури та її сорт. Метою прогнозування є визначення майбутньої температури зерна, що дозволяє оцінити ризики самозігрівання та необхідність коригування

вентиляційних режимів. Прогноз здійснюється як для середньої температури зернової маси, так і для температури в кожній із 200 вимірювальних точок, що дозволяє отримати детальну картину її розподілу всередині сховища.

З метою контролю та управління температурою зерна у сховищах традиційно використовуються статистичні математичні моделі прогнозування, які дозволяють оцінювати майбутні значення температури на основі історичних даних та фізичних закономірностей. До таких методів належать [1]:

- Регресійні моделі, зокрема лінійна регресія, яка використовується для прогнозування температури залежно від середньої зовнішньої температури, вологості повітря та інших факторів [3]. Наприклад, у роботах із прогнозування кліматичних умов зерносховищ застосовується багатофакторна регресія, що враховує зміну температури залежно від вентиляційних параметрів.

- Авто-регресійні методи (ARIMA, SARIMA) широко застосовуються для аналізу часових рядів, що описують динаміку температури зерна в різні періоди року. Ці методи показали ефективність у прогнозуванні змін температури у короткостроковій перспективі, особливо у контрольованих умовах [4].

- Фізичні моделі теплопередачі, які описують баланс тепла всередині зернової маси з урахуванням теплообміну з навколишнім середовищем. Наприклад, для оцінки ефективності вентиляційних режимів застосовуються диференціальні рівняння теплопровідності, що дозволяють визначити, як швидко температура зерна змінюється під впливом зовнішніх факторів.

Однак статистичні підходи мають суттєві обмеження, оскільки не враховують складні нелінійні залежності між різними параметрами, а також не завжди адаптуються до змінних умов зберігання. Децентралізовані підходи до управління температурою в багатозональних системах вже застосовуються у практиці енергоменеджменту та можуть бути адаптовані для прогнозування температури зерна [5]. Наприклад, температурні коливання можуть залежати не лише від зовнішніх умов, а й від внутрішніх процесів у зерновій масі, таких як біологічна активність чи нерівномірний розподіл вологості.

З огляду на це, у сучасних дослідженнях все більше уваги приділяється інтелектуальним методам прогнозування, зокрема нейро-нечітким моделям, які поєднують можливості штучних нейронних мереж та нечіткої логіки. Нейро-нечіткі системи, такі як ANFIS (AdaptiveNeuro-FuzzyInferenceSystem), показують високу ефективність у прогнозуванні складних динамічних процесів завдяки здатності адаптивно коригувати свої параметри на основі накопичених даних [6].

Практичне застосування нейро-нечітких моделей у прогнозуванні температури зерна вже має позитивні результати. Наприклад, у деяких автоматизованих системах моніторингу зерносховищ застосовуються ANFIS-моделі для коригування режимів вентиляції залежно від прогнозованих значень температури та вологості. У таких системах, зібрані сенсорами дані передаються в нейро-нечітку систему, яка аналізує тренди змін температури та формує рекомендації щодо активації охолоджувальних механізмів. Це дозволяє суттєво знизити ризик самозігрівання зерна та оптимізувати енергетичні витрати [7].

Крім того, нейро-нечіткі системи успішно застосовуються у великих зерносховищах із розподіленою системою датчиків, що передають дані в реальному часі. У таких випадках використання статистичних математичних моделей ускладнюється через велику кількість параметрів, тоді як нейро-нечіткі системи здатні ефективно узагальнювати інформацію та знаходити закономірності навіть за наявності шумів у вхідних даних.

Таким чином, метою цього дослідження є порівняння ефективності статистичних методів прогнозування температури зерна та нейро-нечітких моделей, а також визначення їхніх переваг і недоліків у контексті використання в умовах зерносховищ. Основні завдання статті включають:

- аналіз наявних методів прогнозування;

- проведення експериментального дослідження на основі реальних або симуляційних даних;
- оцінку точності прогнозування різними підходами;
- формулювання практичних рекомендацій щодо вибору оптимальної моделі.

Результати дослідження можуть бути корисними для аграрних підприємств, інженерів та науковців, які займаються розробкою систем моніторингу та управління умовами зберігання зерна, а також сприятимуть підвищенню ефективності технологічних рішень у сфері агрологістики.

Датасет

Для навчання та тестування моделей використовувалися реальні дані, які були запозичені у авторів статі «Temperature Forecasting for Stored Grain: A Deep Spatiotemporal Attention Approach». Вони отримані з автоматизованої IoT-системи моніторингу та контролю стану зерна в зерносховищі. Ця система складається з комп'ютерної, мережевої та електронної технологій, що забезпечують виявлення стану зерна, передачу, збереження, аналіз даних та контроль за зерносховищем. Основні компоненти системи включають хост-комп'ютер, інтерфейс передачі даних, допоміжний комп'ютер, сенсори для моніторингу зерна, термометричні кабелі та програмне забезпечення для аналізу стану зберігання [2].

Дані збиралися з мережі з 200 датчиків, із частотою три рази на тиждень у період з 3 липня 2015 року по 17 грудня 2018 року, після чого були інтерпольовані до щоденних значень. Таким чином, загальний обсяг датасету складає приблизно 1500 записів, що дозволяє оцінювати як короткострокові, так і довгострокові тенденції змін температури зерна. Сенсори були організовані у трьох площинах: x - y , x - z та y - z і в більшості випадків залишалися зануреними в зернову масу. Оскільки процеси завантаження та вивантаження зерна можуть призводити до зміни рівня заповненості сховища, частина сенсорів могла опинитися у верхніх шарах або в зоні повітряного простору. Відстань між горизонтальними термометричними кабелями не перевищувала 5 м, між вертикальними — 2 м, а відстань від кабелів до поверхні зерна, дна та стін зерносховища складала 0.3 – 0.5 м. Дані з температурних сенсорів передавалися до центрального керуючого комп'ютера для подальшого аналізу [2].

Окрім реальних даних, використовувалися синтетичні дані, змодельовані на основі рівнянь теплопровідності зернової маси. Для перевірки коректності цих даних їх порівнювали з реальними спостереженнями. Основним завданням аналізу було використання історичних температурних даних та відповідних метеорологічних параметрів для прогнозування температури зерна на майбутні t днів за допомогою просторово-часових моделей.

Прогнозування температурного розподілу всередині сховища розглядалося для кожної окремої вимірювальної точки, що дозволяє оцінити локальні зміни температури та потенційні зони перегріву.

Маса зерна у сховищі не вимірювалася безпосередньо, але її вплив враховувався через щільність насипу та теплопровідність зернової маси, які визначають особливості розподілу температури. Оскільки щільність насипу впливає на теплообмінні процеси та швидкість розповсюдження тепла, цей параметр використовувався у моделях прогнозування як один із ключових факторів.

Температура навколишнього середовища – один із ключових факторів, що впливає на температуру зерна, особливо у сховищах з природною або недостатньою контрольованою вентиляцією. Вплив зовнішньої температури залежить від теплоізоляції сховища та особливостей теплопереносу в зерновій масі. Вологість повітря – визначає інтенсивність теплообміну між зерном і навколишнім середовищем, а також впливає на можливі процеси самозігрівання. Вологість зерна – безпосередньо впливає на швидкість дихання зерна та ризик розвитку мікрофлори, що може спричинити підвищення температури. Щільність

насипу – характеризує ступінь ущільнення зернової маси, що може впливати на рівномірність температурного розподілу. Рівень вентиляції – визначає ефективність охолодження зерна. Вентиляція створює штучний повітряний потік у зерносховищі, який забезпечує зниження температури та вологості зерна. Важливо відзначити, що природний рух повітря без активної вентиляції є мінімальним через щільність насипу. Початкова температура зерна – визначає вихідні умови прогнозування, оскільки зерно може надходити у сховище з різними параметрами залежно від умов збору та транспортування. Час зберігання – параметр, що дозволяє оцінити динаміку змін температури зерна у процесі його зберігання. Тип зернової культури та її сорт.

У моделі прогнозування культура та сорт зерна враховуються через їхній вплив на теплофізичні властивості зернової маси. Зокрема, різні культури (пшениця, кукурудза, ячмінь тощо) мають відмінні показники щільності насипу, теплопровідності та вологості, що впливає на розподіл температури та процеси теплообміну. Для цього використовувалися середньостатистичні значення теплофізичних характеристик кожної культури, отримані з наукових джерел та експериментальних даних. Крім того, сорт зерна визначає рівень його вологості та стійкість до процесів самозігрівання, що також було враховано у моделі як додатковий параметр.

Перед використанням набору даних було проведено валідацію – видалення або заміна аномальних значень, які могли виникнути через помилки вимірювання або збої в системі збору даних. Також було проведено нормалізацію параметрів – приведення значень до однакового масштабу для коректного використання в нейро-нечітких моделях.

Набір даних було розподілено на навчальну вибірку (70 %), валідаційну (15 %) та тестову (15 %) для належного навчання моделей та оцінки їхньої точності.

Для навчання моделей використовувалися часові ряди температури зерна, а також метеорологічні параметри. Використовувалася методика ковзного вікна (time window), що дозволяє прогнозувати майбутнє значення на основі попередніх спостережень. Розмір вікна складав 14 днів, тобто модель отримувала на вхід значення температури за останні 14 днів та прогнозувала температуру на наступний день. Цей підхід застосовувався як для даних температури зерна з 200 датчиків, так і для зовнішніх метеорологічних параметрів (температура навколишнього середовища, вологість повітря, рівень вентиляції тощо).

Щоб забезпечити коректне порівняння методів прогнозування, всі моделі використовували однакові вхідні параметри та навчальні набори даних. Датасет, на якому навчалися нейронні мережі, містив ті ж самі параметри, що й вхідні змінні для нечіткої логіки та статистичних методів. Температура зерна як цільова змінна визначалася однаково для всіх методів. Крім того, тестування всіх моделей відбувалося на спільному тестовому наборі, що гарантувало рівні умови порівняння. Це виключало можливість того, що моделі прогнозують різні процеси, наприклад, коли нейронна мережа навчається на одному наборі даних, а нечітка логіка – на іншому. Завдяки цьому результати порівняння залишаються об'єктивними та статистично коректними.

Методи та технології

Статистичні підходи широко застосовуються для прогнозування температури зерна, оскільки вони дозволяють аналізувати закономірності змін температури на основі історичних даних. Основні методи включають:

- Лінійна та поліноміальна регресія – використовуються для встановлення зв'язку між температурою зерна та факторами, такими як температура навколишнього середовища, вологість повітря та вентиляція. Поліноміальна регресія дозволяє враховувати нелінійні залежності, але при високих ступенях полінома може призводити до перенавчання [3].

- ARIMA (AutoRegressive Integrated Moving Average) – популярний метод аналізу часових рядів, що враховує попередні значення температури, тренди та випадкові коливання.

Ефективний для короткострокового прогнозування, але потребує ретельного налаштування параметрів [4].

- Експоненційне згладжування (методи Хольта, Вінтерса) – надає більшу вагу останнім спостереженням, що дозволяє швидше реагувати на зміни температури. Використовується для прогнозування тенденцій, зокрема з урахуванням сезонних коливань [8].

Статистичні методи є менш ресурсозатратними та добре працюють у короткостроковій перспективі, але їх точність знижується при складних нелінійних взаємозв'язках між параметрами [9].

Інтелектуальні методи, що поєднують можливості штучних нейронних мереж та нечіткої логіки, дозволяють будувати більш гнучкі та адаптивні моделі прогнозування. Основні підходи:

- ANFIS (Adaptive Neuro-Fuzzy Inference System) – нейро-нечітка модель, що поєднує штучні нейронні мережі з нечіткою логікою та базується на нечіткій системі виводу Sugeno. Завдяки адаптивному налаштуванню параметрів на основі навчальних даних, ANFIS демонструє високу точність навіть у довгостроковому прогнозуванні, хоча потребує значних обчислювальних ресурсів [6, 10, 11].

- Mamdani – класичний підхід нечіткої логіки, який дозволяє враховувати складні взаємозв'язки між параметрами. Забезпечує інтуїтивно зрозумілі результати, але має високу обчислювальну складність [12, 13].

Оскільки процес зміни температури зерна у сховищі є нелінійним і залежить від багатьох взаємопов'язаних факторів, доцільним є використання нейро-нечітких систем, які поєднують можливості штучних нейронних мереж та нечіткої логіки. Раніше такі підходи ефективно застосовувалися для координації температурного контролю у багатозональних об'єктах, що підтверджує їхню придатність для прогнозування зернових мас [5]. У рамках дослідження було обрано такі моделі як ANFIS та Mamdani. Процес фазифікації для обох моделей є однаковим і передбачає визначення нечітких множин для вхідних змінних. Основна різниця між цими моделями полягає в методі дефазифікації та побудові правил виводу.

Модель Mamdani базується на нечітких правилах типу «Якщо – то» та дозволяє враховувати складні взаємозв'язки між параметрами, включаючи як кількісні, так і якісні змінні [13]. База знань формувалася на основі комбінації експертних правил та статистичного аналізу зібраних даних. Загальний обсяг бази знань складав понад 50 правил, що визначали взаємозв'язки між температурою зерна, рівнем вентиляції, вологістю та іншими параметрами.

Наприклад:

- «Якщо температура навколишнього середовища висока і вентиляція низька, то температура зерна підвищується.»

- «Якщо зерно – кукурудза, а рівень вентиляції низький, то температура зерна підвищується швидше, ніж у пшениці.»

Для кожного з вхідних параметрів визначалися нечіткі підмножини. Наприклад, температура навколишнього середовища, початкова температура зерна та рівень вентиляції поділялися на низький, середній та високий рівень, тоді як вологість повітря та зерна класифікувалися як низька, середня або висока. Для якісних змінних використовувалися категорійні нечіткі множини: для типу зернової культури було визначено, що пшениця має нижчу теплопровідність порівняно з кукурудзою, а її вологість повільніше змінюється у процесі зберігання. База правил будувалася на основі експертних знань і статистичного аналізу реальних даних. Дефазифікація здійснювалася методом центру ваги, що дозволило отримати чітке значення прогнозованої температури [14].

ANFIS (Adaptive Neuro-Fuzzy Inference System) – нейро-нечітка модель, яка базується на нечіткій системі виводу Sugeno. Вона поєднує можливості штучних нейронних мереж та нечіткої логіки, що дозволяє адаптивно налаштовувати параметри членських функцій на основі навчальних даних [15]. Процедура фазифікації вхідних змінних для ANFIS (Sugeno) і

Mamdani є однаковою, передбачаючи перетворення чітких значень у нечіткі множини. Основна різниця між цими моделями полягає в методі дефазифікації та способі формування правил виводу. У випадку ANFIS система автоматично оптимізує вагові коефіцієнти та структуру нечітких правил у процесі навчання [11]. Спочатку система генерувала базові нечіткі правила, такі як:

- «Якщо вологість зерна висока і щільність насипу велика, то температура зерна зростає повільно.»
- «Якщо зерно – пшениця і рівень вентиляції низький, то температура зерна змінюється відповідно до функції $f(\text{вологість, температура середовища})$ »
- «Якщо температура навколишнього середовища висока і рівень вентиляції низький, то температура зерна $= 0.8 \times \text{температура середовища} + 0.2 \times \text{рівень вентиляції}$ »
- «Якщо вологість зерна висока і щільність насипу велика, то температура зерна $= 0.5 \times \text{вологість} + 0.3 \times \text{щільність насипу}$ »

На другому етапі виконувалася оптимізація цих правил за допомогою алгоритмів зворотного поширення помилки та методу найменших квадратів. Для фазифікації використовувалися гаусові функції належності, а висновок визначався лінійною комбінацією вхідних змінних. Такий підхід забезпечував високу точність прогнозування за рахунок здатності моделі адаптуватися до даних [10].

Таким чином, кожна із моделей має свої переваги залежно від цільового використання. Mamdani є ефективною для експертних систем, де важливо інтуїтивне пояснення процесів, але вона потребує значних обчислювальних ресурсів. ANFIS відзначається високою точністю прогнозування завдяки адаптивному навчанню та оптимізації параметрів. Використання цих підходів дозволяє підвищити ефективність прогнозування температури зерна та оптимізувати процес його зберігання [15].

Для реалізації та аналізу моделей було використано кілька програмних середовищ та бібліотек. Основною мовою програмування для розробки та тестування моделей був Python. Для побудови та навчання нейронних мереж, зокрема нейро-нечітких моделей ANFIS та Mamdani, застосовували TensorFlow та Keras. Статистичні методи прогнозування, такі як лінійна регресія, поліноміальна регресія та ARIMA, реалізовували за допомогою бібліотеки Scikit-learn. Часові ряди, зокрема ARIMA та експоненційне згладжування, обробляли за допомогою Statsmodels. Для реалізації нечітких моделей Mamdani та ANFIS (Sugeno) використовували Fuzzy Logic Toolbox у MATLAB. Обробку, аналіз та нормалізацію вхідних даних виконували за допомогою Pandas і NumPy, а для візуалізації отриманих результатів, графіків прогнозів та порівняння моделей застосовували Matplotlib і Seaborn. Початкову обробку даних та верифікацію статистичних характеристик здійснювали у MS Excel. Навчання моделей проводилося у локальному середовищі Python, а тестування деяких методів виконували у MATLAB.

Для об'єктивного порівняння ефективності статистичних методів прогнозування та нейро-нечітких моделей було використано три основні метрики оцінки точності: середньоквадратичну похибку (MSE), середню абсолютну похибку (MAE) та корінь середньоквадратичної похибки (RMSE). Вибір цих метрик обумовлений їхньою здатністю характеризувати точність прогнозування та рівень відповідності моделей реальним даним.

Середньоквадратична похибка (MeanSquaredError, MSE) є однією з найпоширеніших метрик оцінки прогнозів. Вона розраховується за формулою:

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

де y_i – фактичне значення температури зерна; \hat{y}_i – прогнозоване значення; n – кількість спостережень.

MSE дозволяє оцінити середній рівень помилки моделі в квадратних одиницях вимірювання, що робить її чутливою до великих відхилень. Високе значення MSE свідчить

про низьку точність прогнозу, а менше значення вказує на кращу відповідність реальним даним.

У прогнозуванні температури зерна MSE може бути використана для порівняння різних моделей на основі їх здатності мінімізувати значні помилки, які можуть спричинити перегрівання зернової маси та зниження її якості.

Середня абсолютна похибка (MeanAbsoluteError, MAE) визначає середнє значення абсолютних відхилень прогнозованих значень від реальних і розраховується за формулою:

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

На відміну від MSE, ця метрика не враховує квадратичні відхилення, а тому менш чутлива до аномальних значень. Вона дозволяє оцінити середню похибку прогнозу в тих же одиницях вимірювання, що й прогнозована величина (°C у нашому випадку).

MAE корисна для аналізу середнього рівня похибки прогнозування температури зерна, оскільки дозволяє зрозуміти, наскільки в середньому прогноз відрізняється від фактичних значень.

Корінь середньоквадратичної похибки (RootMeanSquaredError, RMSE) є ще однією важливою метрикою оцінки точності прогнозів, яка розраховується як квадратний корінь із MSE:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (3)$$

RMSE має ті ж одиниці вимірювання, що й прогнозована величина, що робить її зручною для інтерпретації. Вона є чутливою до великих відхилень, оскільки підносить помилки до квадрата перед усередненням.

RMSE використовується для оцінки точності прогнозу, особливо коли важливо враховувати великі відхилення. У контексті прогнозування температури зерна низьке значення RMSE означає кращу відповідність прогнозованих значень реальним спостереженням.

Використання цих трьох метрик у дослідженні дозволяє отримати всебічну оцінку точності моделей прогнозування температури зерна у сховищах. MSE та MAE дозволяють кількісно оцінити рівень помилки, а RMSE дає змогу оцінити, наскільки великі відхилення впливають на прогноз.

Результати моделювання

Після проведення тестування статистичних методів прогнозування та нейро-нечітких моделей було виконано порівняльний аналіз їхньої точності. Оцінка базувалася на використанні метрик MSE, MAE та RMSE, які дозволяють визначити рівень відхилення прогнозованих значень від фактичних.

Середньоквадратична похибка характеризує середній рівень помилки в квадратних одиницях вимірювання. Результати MSE для всіх досліджених моделей наведено в таблиці 1.

Таблиця 1

Значення середньоквадратичної похибки (MSE) для різних моделей прогнозування

Модель	Горизонт 3 дні	Горизонт 7 днів	Горизонт 14 днів
Лінійна регресія	1,02	1,45	2,31
Поліноміальна регресія	0,85	1,22	2,05
ARIMA	0,78	1,18	1,92
Експоненційне згладжування	0,91	1,30	2,25
ANFIS(Sugeno)	0,52	0,79	1,28
Mamdani	0,60	0,91	1,42

Як видно з таблиці, статистичні методи (лінійна регресія та ARIMA) демонструють прийнятну точність при короткостроковому прогнозуванні (до 7 днів), проте їх похибка зростає при збільшенні горизонту прогнозування. Нейро-нечіткі моделі, зокрема ANFIS, демонструють значно менші значення MSE, що підтверджує їхню здатність до моделювання складних нелінійних залежностей.

MAE оцінює середню різницю між прогнозованими та фактичними значеннями температури зерна.

Таблиця 2

Значення середньої абсолютної похибки (MAE) для різних моделей прогнозування

Модель	Горизонт 3 дні	Горизонт 7 днів	Горизонт 14 днів
Лінійна регресія	0,85	1,15	1,89
Поліноміальна регресія	0,72	1,05	1,75
ARIMA	0,66	0,98	1,65
Експоненційне згладжування	0,78	1,10	1,82
ANFIS(Sugeno)	0,41	0,64	1,02
Mamdani	0,50	0,75	1,18

Результати MAE підтверджують, що нейро-нечіткі моделі забезпечують менший рівень похибки порівняно з статистичними методами. Найнижче середнє абсолютне відхилення демонструє ANFIS, що робить цю модель оптимальним вибором для прогнозування температури зерна у довгостроковій перспективі.

Коренева середньоквадратична похибка (RMSE) є важливим показником точності прогнозування, який враховує як зміщення, так і дисперсію похибки. Менше значення RMSE вказує на кращу відповідність моделі реальним даним.

Таблиця 3

Значення кореня середньоквадратичної похибки (RMSE) для різних моделей прогнозування

Модель	Горизонт 3 дні	Горизонт 7 днів	Горизонт 14 днів
Лінійна регресія	1,01	1,20	1,52
Поліноміальна регресія	0,92	1,10	1,43
ARIMA	0,88	1,09	1,39
Експоненційне згладжування	0,95	1,14	1,50
ANFIS(Sugeno)	0,72	0,89	1,13
Mamdani	0,77	0,95	1,19

Як показують результати, статистичні методи, такі як лінійна та поліноміальна регресії, демонструють вищі значення RMSE порівняно з іншими підходами, що свідчить про їхню нижчу точність. Наприклад, для 14-денного прогнозу RMSE лінійної регресії становить 1,52, а для поліноміальної – 1,43. Метод ARIMA показав дещо кращі результати (RMSE = 1,39 для горизонту 14 днів), однак також поступається нейро-нечітким моделям.

Проте значення цих метрик потрібно інтерпретувати з урахуванням діапазону температур у зерносховищі. У дослідженні температура зерна змінюється в межах від 8°C до 15°C, тобто максимальна різниця становить 7°C. Таким чином, наприклад, RMSE = 1,01°C означає відносну похибку близько 14,4 % від усього діапазону температур. Це слід враховувати при порівнянні моделей, оскільки однакове значення RMSE може мати різний вплив залежно від ширини діапазону температур.

Результати дослідження підтверджують перевагу нейро-нечітких моделей над статистичними методами прогнозування температури зерна. Зокрема, модель ANFIS, забезпечила найнижчі значення RMSE для всіх горизонтів прогнозування (0,72, 0,89, 1,13 відповідно), що свідчить про її високу прогностичну здатність. Інша нечітка модель, Mamdani, також показала конкурентоспроможні результати, суттєво перевершуючи

статистичні підходи. Це підтверджує ефективність застосування нейро-нечітких систем для прогнозування температури зерна у сховищах.

Одним із ключових аспектів при виборі моделі прогнозування є не лише її точність, але й продуктивність та обчислювальна складність. Високоточні моделі можуть вимагати значних обчислювальних ресурсів, що може бути критичним фактором при їх практичному використанні.

Час, необхідний для навчання моделі, та швидкість виконання прогнозу є важливими критеріями для практичного застосування. У таблиці 4 наведено середній час навчання та виконання прогнозу для кожної моделі.

Таблиця 4

Час навчання та прогнозування для різних моделей

Модель	Час навчання (с)	Час прогнозу на 1000 точок (мс)
Лінійна регресія	0,5	1,2
Поліноміальна регресія	0,8	1,5
ARIMA	32	8,9
Експоненційне згладжування	1,5	2,3
ANFIS(Sugeno)	450	12,5
Mamdani	505	15,7

Як видно з таблиці 4, найменший час навчання та прогнозування демонструють статистичні методи, такі як лінійна та поліноміальна регресія, а також експоненційне згладжування. Моделі ARIMA потребують значно більше часу на навчання, особливо за великого обсягу даних.

З іншого боку, нейро-нечіткі моделі (ANFIS та Mamdani) вимагають значно більше часу на навчання, що пояснюється складністю їхньої структури, необхідністю оптимізації параметрів нечіткої логіки та нейронної мережі. Проте, після навчання, час прогнозування залишається в межах прийняттого рівня.

Обчислювальна складність моделей визначається кількістю операцій, необхідних для навчання та прогнозування. Нижче наведено загальні оцінки складності досліджуваних моделей:

- Лінійна регресія: $O(n)$ — проста модель із лінійною складністю для навчання і прогнозування;
- Поліноміальна регресія: $O(n^d)$, де d — степінь полінома, що може значно ускладнити обчислення при високих значеннях d ;
- ARIMA: $O(n^2)$ — швидка і обчислювально ефективна модель як для навчання, так і для прогнозування;
- Експоненційне згладжування: $O(n)$ — швидка обчислювально ефективна модель;
- ANFIS: $O(nm)$ для навчання, оскільки параметри нечітких правил і функцій належності оптимізуються в процесі тренування; прогнозування після навчання виконується значно швидше – $O(m)$ [16];
- Mamdani: $O(nm^2)$ — для навчання, оскільки ця модель має об'ємну базу правил і потребує додаткових обчислень для дефазифікації вихідних значень; прогнозування також є ресурсоємним через значну кількість логічних операцій[14, 13].

Таким чином, статистичні методи мають нижчу обчислювальну складність, що робить їх придатними для роботи з великими наборами даних у режимі реального часу. Нейро-нечіткі моделі забезпечують високу точність, проте вимагають значних обчислювальних ресурсів, що може бути обмеженням для практичного використання у великих зерносховищах із постійним моніторингом температури.

Аналіз результатів

Отримані результати показують, що RMSE для нейро-нечітких моделей становить 0,72 – 1,13 залежно від горизонту прогнозування, тоді як для статистичних методів цей показник є вищим. Враховуючи, що температура зерна змінюється в межах 8 – 15°C (різниця 7°C), це означає, що середнє відхилення прогнозованих значень у моделі ANFIS складає приблизно 10 – 16 % від усього діапазону можливих значень температури. У практичному сенсі це свідчить про достатню точність прогнозу для ефективного управління вентиляційними процесами.

Статистичні методи, такі як ARIMA, показали прийнятну точність для короткострокового прогнозування (3 – 7 днів), однак на довгих горизонтах (14 днів) значно поступалися нейро-нечітким моделям. Отримані результати підтверджують доцільність застосування нейро-нечітких моделей у системах моніторингу зерносховищ, що дозволяє підвищити ефективність управління процесом зберігання зерна.

Нейро-нечіткі моделі ANFIS та Mamdani вимагають більше часу на навчання, але їх перевага полягає в здатності моделювати складні взаємозв'язки між параметрами, що робить їх ефективнішими для довгострокового прогнозування температури зерна.

Вибір оптимального методу прогнозування залежить від цілей дослідження. Якщо необхідно швидке прогнозування у короткостроковій перспективі, доцільно використовувати ARIMA або експоненційне згладжування. Для довгострокового прогнозування або в умовах значної мінливості даних нейро-нечіткі моделі є більш ефективними.

Висновки

У цій статті проведено аналіз точності прогнозування температури зерна в сховищах, порівнюючи статистичні методи та нейро-нечіткі моделі. Дослідження підтвердило, що нейро-нечіткі системи, зокрема ANFIS, що базується на нечіткій системі виводу Sugeno, та Mamdani, демонструють найвищу точність завдяки здатності ефективно моделювати складні нелінійні залежності. Це робить їх особливо перспективними для використання в умовах, де потрібен детальний аналіз та висока прогностична здатність. База знань нечітких моделей формувалася на основі експертних правил і статистичного аналізу зібраних даних, що дозволило забезпечити коректність нечітких висновків. Водночас їх застосування пов'язане з підвищеними вимогами до обчислювальних ресурсів і необхідністю наявності великого обсягу навчальних даних, що може ускладнювати їх впровадження в деяких практичних сценаріях.

Методи ARIMA та експоненційного згладжування показали високу ефективність у середньостроковому прогнозуванні, пропонуючи оптимальний баланс між точністю та швидкістю обчислень. Вони можуть бути ефективними для випадків, коли важливим є швидке отримання результатів без значних обчислювальних витрат. Поліноміальна регресія також продемонструвала конкурентоспроможні результати, особливо при помірному обсязі вхідних даних, що свідчить про її потенційну корисність у випадках, коли нейро-нечіткі моделі є надмірно складними для використання.

Лінійна регресія виявилася найменш точною з усіх розглянутих методів, однак вона залишається корисним інструментом для швидкої оцінки загальних тенденцій і може бути застосована в ситуаціях, де висока точність прогнозування не є критичною.

Результати аналізу підтверджують, що вибір методу прогнозування значною мірою залежить від умов використання. Якщо основною метою є отримання максимально точної оцінки, доцільно застосовувати нейро-нечіткі моделі, особливо в автоматизованих системах моніторингу зерносховищ. У випадках, коли необхідний компроміс між точністю та швидкістю розрахунків, ефективним рішенням є ARIMA або експоненційне згладжування. Якщо ж система працює в умовах обмежених обчислювальних ресурсів, найкращими варіантами будуть експоненційне згладжування або лінійна регресія, які не потребують значних обчислювальних потужностей.

Отримані результати підтверджують доцільність застосування нейро-нечітких підходів у системах моніторингу температури зерна, оскільки вони забезпечують точніший прогноз і дозволяють зменшити ризик самозігрівання та втрат зернової маси. Подальші дослідження можуть бути зосереджені на розробці гібридних моделей, що поєднують сильні сторони статистичних та нейро-нечітких методів, а також на адаптації прогнозних систем до специфічних умов зберігання зерна.

СПИСОК ЛІТЕРАТУРИ

1. Grain Pile Temperature Forecasting from Weather Factors: A Support Vector Regression Approach / S. Duan et al. *2019 IEEE/CIC International Conference on Communications in China (ICCC)*, Changchun, China, 2019. P. 255–260.
2. Temperature Forecasting for Stored Grain: A Deep Spatiotemporal Attention Approach / S. Duan et al. *IEEE Internet of Things Journal*. Dec. 1, 2021. Vol. 8, № 23. P. 17147–17160.
3. Marasco S., Marano G. C., Cimellaro G. P. Evolutionary polynomial regression algorithm combined with robust Bayesian regression. *Advances in Engineering Software*. 2022. Vol. 167. P. 103101. ISSN 0965-9978.
4. Siami-Namini S., Tavakoli N., Siami Namin A. A Comparison of ARIMA and LSTM in Forecasting Time Series. *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*. Orlando, FL, USA, 2018. P. 1394–1401. DOI: 10.1109/ICMLA.2018.00227.
5. Yukhymchuk M., Dubovoi V., Kovtun V. Decentralized coordination of temperature control in multiarea premises. *Complexity*. 2022. Vol. 2022, № 1. Art. ID 2588364. P. 1 – 18. <https://doi.org/10.1155/2022/2588364>
6. Tektas M. Weather Forecasting Using ANFIS and ARIMA Models: Case Study for İstanbul. *Environmental Research Engineering and Management*. 2010. Vol. 51. P. 5–10. DOI: 10.5755/j01.ere.m.51.1.58.
7. Deep Spatio-Temporal Attention Model for Grain Storage Temperature Forecasting / S. Duan et al. *2020 IEEE 26th International Conference on Parallel and Distributed Systems (ICPADS)*. Hong Kong, 2020. P. 593–600.
8. Rendon-Sanchez J. F., de Menezes L. M. Structural combination of season alex potentials smoothing forecasts applied to load forecasting. *European Journal of Operational Research*. 2019. Vol. 275, № 3. P. 916–924. ISSN 0377-2217.
9. Ali Z., Bhaskar S. B. Basic statistical tools in research and data analysis. *Indian Journal of Anaesthesia*. 2016. Vol. 60, № 9. P. 662–669.
10. Bhardwaj R., Duhoon V. Real Time Prediction of Temperature using ANFIS-SUGENO Model. *International Journal of Engineering and Advanced Technology*. 2019. №9 (1). P. 461–469. URL: <https://www.ijeat.org/wp-content/uploads/papers/v9i1/A9555109119.pdf>.
11. Alves K. S. T. R., de Jesus C. D., de Aguiar E. P. A new Takagi–Sugeno–Kang model for time series forecasting. *Engineering Applications of Artificial Intelligence*. 2024. Vol. 133, Part C. P. 108155. ISSN 0952-1976.
12. Fuzzy logic model for temperature prediction / A. M. Harahap et al. *Journal of Physics: Conference Series*. 2021. Vol. 1783. P. 012039.
13. Ardi Y., Effendi S., Nababan E. B. Mamdani and Sugeno Fuzzy Performance Analysis on Rain fall Prediction. *Rand wick International of Social Science Journal*. 2021. Vol. 2, № 2. P. 161–177.
14. Yang K., Zhao L. Load Forecasting Model Based on Amendment of Mamdani Fuzzy System. 2009 *5th International Conference on Wire less Communications, Networking and Mobile Computing*. Beijing, China, 2009. P. 1–4.
15. Short-term air temperature prediction by adaptive en euro-fuzzy inference system (ANFIS) and long short-term memory (LSTM) network / A. Sekertekin et al. *Meteorology and Atmospheric Physics*. 2021. Vol. 133. P. 943–959.
16. Use of Artificial Neuro-Fuzzy Network Takagi-Sugeno-Kangto the Assessment of Technical Condition of Construction Objects / O. Komandyrov et al. *Management of Development of Complex Systems*. 2020. P. 107–112. DOI: 10.32347/2412-9933.2020.42.107-112.

Стаття надійшла до редакції 05.03.2025.

Стаття пройшла рецензування 17.03.2025.

Ліщук Андрій Романович – аспірант кафедри комп'ютерних систем управління, e-mail: a.r.lishchuk@gmail.com.

Вінницький національний технічний університет.