

Дармосюк В. М.,
канд. фіз.-мат. наук, доцент,
доцент кафедри вищої та прикладної математики
Миколаївський національний аграрний університет, м. Миколаїв

МОДЕЛІ ARIMA ДЛЯ ПРОГНОЗУВАННЯ ЧАСОВИХ РЯДІВ У СІЛЬСЬКОМУ ГОСПОДАРСТВІ

Сільське господарство є ключовим сектором для будь-якої економіки, і точне прогнозування сільськогосподарського виробництва відіграє життєво важливу роль у забезпеченні продовольчої безпеки, управлінні ланцюгами поставок та прийнятті політичних рішень. Методи аналізу та прогнозування часових рядів набувають все більшого значення в сільському господарстві через наявність великих масивів даних та потребу в надійних прогнозах. Серед різних методів прогнозування часових рядів моделі авторегресійного інтегрованого ковзного середнього (ARIMA) привертають значну увагу завдяки своїй здатності фіксувати та моделювати складні закономірності в даних часових рядів.

ARIMA-моделі – це клас статистичних моделей, які описують і прогнозують дані часових рядів на основі їхніх власних минулих значень і минулих помилок. Ці моделі особливо корисні, коли основний процес, що генерує часовий ряд, невідомий або важко піддається явному моделюванню. Моделі ARIMA поєднують три компоненти: авторегресійний (AR) компонент, який враховує вплив минулих значень на поточне значення; інтегрований (I) компонент, який враховує нестационарність даних шляхом диференціювання; і компонент ковзного середнього (MA), який включає залежність між спостереженням і залишковою похибкою від моделі ковзного середнього, застосованої до запізнілих спостережень. Загальний запис моделі ARIMA має вигляд $ARIMA(p, d, q)$, де p – порядок компонента авторегресії; d – ступінь диференціювання, необхідний для того, щоб зробити часовий ряд стаціонарним; q – порядок компонента ковзного середнього. Моделі ARIMA успішно застосовуються до різних часових рядів у сільському господарстві для прогнозування. Розглянемо більш детально деякі приклади.

Моделі ARIMA можна використовувати для прогнозування врожайності сільськогосподарських культур на основі історичних даних про врожайність. Ця інформація має вирішальне значення для фермерів, агробізнесу та політиків для планування виробництва, управління ресурсами та забезпечення продовольчої безпеки. Наприклад, M. Amin, M. Amanullah and A. Akbar у своєму дослідженні розробили моделі часових рядів та визначили найкращу модель для прогнозування виробництва пшениці в Пакистані. Були використані великі часові періоди, тобто дані за 1902-2005 роки. Різні моделі часових рядів були підігнані до цих даних за допомогою двох програмних продуктів – JMP та Statgraphics. Дослідники виявили, що найкращою моделлю є ARIMA (1,2,2). На основі цієї обраної моделі вони виявили, що виробництво пшениці в Пакистані становитиме 26623,5 тис. т. у 2020 році та подвоїться у 2060 році порівняно з 2010 роком [1].

Прогнозування поголів'я худоби має важливе значення для управління запасами кормів, розробки селекційних програм та забезпечення сталих виробничих практик. Моделі ARIMA застосовуються для прогнозування поголів'я великої рогатої худоби, овець та птиці на основі історичних даних. Наприклад, Muhammad Qasim та інші у своєму дослідженні визначали відповідну емпіричну модель для прогнозування популяції буйволів у Пакистані, щоб оцінити її майбутній тренд до 2030 року. Вони застосовували різні авторегресійні інтегровані ковзні середні (ARIMA) для популяції буйволів на основі п'ятдесятирічних часових рядів даних. Використовували різні критерії відбору моделей для перевірки надійності моделей ARIMA. На основі цих критеріїв прийшли до висновку, що ARIMA (1,0,0) є більш підходящою моделлю. Крім того, вони також перевіряли припущення підібраної моделі, такі як нормальність та незалежність, щоб з'ясувати більш точні прогнозовані значення. Це дослідження показало, що популяція буйволів, як очікується, збільшиться на 30% до 2030 року за умови, що впродовж прогнозованих років не спостерігатиметься ніяких нерегулярних тенденцій [2].

Точне прогнозування цін на сільськогосподарську продукцію є критично важливим для фермерів, трейдерів та політиків для прийняття обґрунтованих рішень щодо виробничої, маркетингової та торговельної політики. Моделі ARIMA використовуються для прогнозування цін на різні товари, такі як зернові, фрукти та овочі. Наприклад, Kumar Ravi Ranjan та Vaishya Moumita у своєму дослідженні відібрали найкращі моделі шляхом порівняння інформаційних критеріїв Акаїке (AIC), Байєсівського інформаційного критерію (BIC), середньої абсолютної відсоткової помилки (MAPE) та середньоквадратичної помилки (RMSE). Дослідження показало, що ARIMA (1,1,2) та ряд інших виявилися найкраще пристосованими моделями для прогнозування цін на картоплю для штатів Уттер-Прадеш, Західна Бенгалія, Мадхья-Прадеш, Гуджарат, Пенджаб, Тріпура в Індії [3].

Погодні та кліматичні умови суттєво впливають на сільськогосподарське виробництво. Моделі ARIMA можна використовувати для прогнозування метеорологічних змінних, таких як температура, опади та вологість, які є важливими вхідними даними для моделювання врожайності та процесів прийняття рішень. Наприклад, Dimgi T. та інші у своїй статті описали дослідження часових рядів та сезонний аналіз середньомісячних мінімальних і максимальних температур та опадів для басейну річки Бхагіратхі, розташованої в штаті Уттаракханд, Індія. Використані дані за 1901-2000 роки (100 років). Було використано сезонну модель ARIMA і зроблено прогноз на наступні 20 років (2001-2020). Було виявлено, що найбільш придатною моделлю для аналізу часових рядів опадів є ARIMA(0,1,1), а для аналізу температурних даних – ARIMA(0,1,0). Результати прогнозування моделі показують, що прогнозні дані добре узгоджуються з тенденцією даних. Інформація про закономірності та тенденції може допомогти в якості інструменту прогнозування для розробки кращих практик управління водними ресурсами в регіоні [4].

Моделі ARIMA мають кілька переваг у прогнозуванні сільськогосподарських часових рядів:

- моделі ARIMA можуть охоплювати широкий спектр моделей часових рядів, включаючи тренди, сезонність та циклічні компоненти, що робить їх придатними для різних сільськогосподарських застосувань.

- моделі ARIMA ґрунтуються на даних, тобто вони не вимагають явних припущень про основний процес, що генерує часовий ряд. Це особливо корисно, коли основний процес невідомий або важко піддається моделюванню.

- ARIMA-моделі можна застосовувати до одновимірних часових рядів, що спрощує процес моделювання та зменшує потребу в додаткових пояснювальних змінних.

Однак моделі ARIMA також мають певні обмеження, наприклад, припускають, що часовий ряд є стаціонарним або може бути стаціонарним за допомогою диференціювання, що може не виконуватись для певних сільськогосподарських часових рядів, які демонструють нелінійні або структурні зміни. Моделі ARIMA в першу чергу зосереджені на прогнозуванні і можуть не давати уявлення про основні причинно-наслідкові механізми або взаємозв'язки між змінними. Хоча моделі ARIMA можуть відображати закономірності в одновимірних часових рядах, вони не враховують вплив екзогенних змінних, таких як погодні умови або економічні фактори, які можуть бути важливими в сільському господарстві. Щоб подолати деякі з цих обмежень, дослідники вивчали розширення та варіації ARIMA моделей, такі як сезонні ARIMA моделі (SARIMA) для врахування сезонності та ARIMA моделі з екзогенними змінними (ARIMAX) для включення зовнішніх факторів.

ARIMA-моделі виявилися цінним інструментом для прогнозування часових рядів у сільському господарстві. Їх здатність охоплювати складні закономірності та надавати точні прогнози зробила їх широко застосовуваними в таких сферах, як прогнозування врожайності сільськогосподарських культур, прогнозування поголів'я худоби, прогнозування цін на сировинні товари, а також прогнозування погоди та клімату. Важливо визнати обмеження моделей ARIMA і за необхідності вивчати альтернативні або додаткові підходи. Майбутні дослідження повинні зосередитися на інтеграції моделей ARIMA з іншими методами, такими як алгоритми машинного навчання та причинно-наслідкові моделі, щоб підвищити точність прогнозування та інтерпретованість в аграрному контексті.

Список використаних джерел

1. Amin, M., M. Amanullah, and A. Akbar. Time series modeling for forecasting wheat production of Pakistan. *JAPS: Journal of Animal & Plant Sciences*, 24(5): 2014, P. 1444-1451.

2. Qasim, Muhammad, et al. Forecasting buffalo population of Pakistan using autoregressive integrated moving average (ARIMA) time series models. *Proceedings of the Pakistan Academy of Sciences: A. Physical and Computational Sciences* 56.3. 2019. P. 27-36.

3. Kumar Ravi Ranjan, and Moumita Baishya. Forecasting of Potato prices in India: An application of ARIMA model. *Economic Affairs*. 65.4. 2020. P: 473-479.

4. Dimri, T., Ahmad, S. & Sharif, M. Time series analysis of climate variables using seasonal ARIMA approach. *J Earth Syst Sci* 129, 149. 2020.