

Прогнозування динаміки розвитку тваринництва за допомогою моделі SARIMAX

Кулик Анатолій Борисович, к.ф.-м.н., доц.
ORCID: 0000-0002-6629-0253
КНЕУ ім. Вадима Гетьмана, Україна

Тваринництво відіграє важливу роль у житті суспільства, виконуючи багато різних функцій, основною з яких є виробництво продуктів харчування, що впливає на забезпечення продовольчої та економічної безпеки країни та її регіонів. Крім того, галузь виконує інші функції, такі як збереження якості довкілля, підтримання ландшафтів, забезпечення зайнятості населення, соціально-економічний розвиток сільської місцевості тощо. У зв'язку з цим комплексний аналіз функцій та визначення ролі тваринництва у розвитку сільських територій є актуальним.

Прогнозування майбутнього аграрного сектору економіки може бути складним завданням, оскільки воно залежить від багатьох факторів, які можуть бути непередбачуваними. Однак, існують підходи та інструменти, які можуть допомогти у прогнозуванні розвитку аграрного сектору. Одним з таких інструментів є прогнозування часових рядів. Це техніка, яка використовує історичні та поточні дані для прогнозування майбутніх значень протягом певного періоду часу або певної точки в майбутньому. Аналізуючи дані, які ми зберігаємо в минулому, ми можемо приймати обґрунтовані рішення, які можуть спрямовувати нашу бізнес-стратегію та допомагати нам зрозуміти майбутні тенденції (Atalan, 2022).

Метою цієї роботи є прогнозування динаміки розвитку поголів'я великої рогатої худоби і корів на найближчі 24 місяці за допомогою моделі SARIMAX, яка застосовується до дослідження часових рядів.

В основі цього дослідження лежить модель Бокса-Дженкінса (Ediger, Akar, 2007) (Autoregressive Integrated Moving Average, ARIMA), яка дозволяє будувати прогнози на основі часових рядів, тобто історичних спостережень. Модель ARIMA – це комбінація моделі авторегресії (AR) і моделі ковзного середнього (MA)

$$\text{AR (p)} \quad y_t = c + \sum_{n=1}^p \alpha_n y_{t-n} + \varepsilon_t, \quad (1)$$

$$\text{MA (q)} \quad \varepsilon_t = \sum_{n=1}^q \theta_n \varepsilon_{t-n}. \quad (2)$$

Для параметризації моделі використовуються три цілих числа: (p, d, q) .

p – кількість членів авторегресії; d – кількість несезонних відмінностей; q – кількість умов ковзного середнього.

SARIMAX (Seasonal Auto-Regressive Integrated Moving Average with eXogenous factors) – це оновлена версія моделі ARIMA. ARIMA включає авторегресійне інтегроване ковзне середнє, тоді як SARIMAX включає сезонні ефекти та екзогенні фактори з авторегресійним і ковзним середнім компонентом у моделі. Таким чином, SARIMAX є сезонним еквівалентом моделі SARIMA та Auto ARIMA.

Мірою оцінки, яка використовується для прогнозування часових рядів в цьому дослідженні є середньоквадратичне відхилення (RMSE) і середня абсолютна помилка у відсотках (MAPE)

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (E_t - F_t)^2}, \quad MAPE\% = \frac{1}{n} \sum_{i=1}^n \frac{|E_t - F_t|}{E_t}, \quad (3)$$

де E_t і F_t фактичні і прогнозовані значення, n – кількість значень.

Для застосування техніки прогнозування часових рядів потрібно перевірити ряди на стаціонарність за допомогою теста Дікі-Фуллера (ADF)). Цей тест виявляє наявність стохастичної поведінки тенденції у часових рядах за допомогою тесту гіпотези.

H_0 : процес нестационарний; H_1 : процес стаціонарний.

Якщо ряд нестационарний, то за допомогою Box Cox Transformation отримуємо ряд, який буде задовольняти умовам стаціонарності. В основі Box Cox Transformation лежить експонента і коефіцієнт λ , який змінюється від -5 до 5. Розглядаються всі значення λ і вибирається значення, яке дає найкраще наближення кривої нормального розподілу. Перетворення Y має вигляд:

$$y(\lambda) = \begin{cases} \frac{y^{\lambda}-1}{\lambda}, & \text{if } \lambda \neq 0; \\ y, & \text{if } \lambda = 0. \end{cases} \quad (4)$$

Розглянемо динаміку змін кількості великої рогатої худоби (в т.ч. корови) в період з 1 січня 2008 року по 1 січня 2023 року в Хмельницькій області (рис. 1).

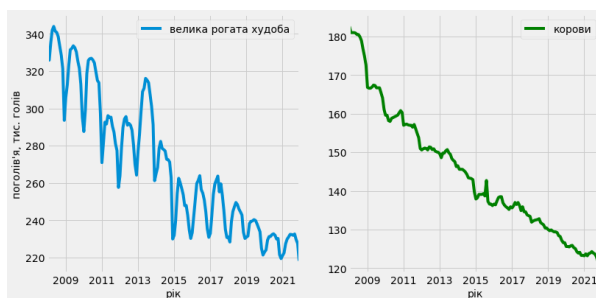


Рисунок 1 – Динаміка зміни поголів'я великої рогатої худоби і корів (тис. голів) з 1 січня 2008 року по 1 січня 2023.

Джерело: розроблено автором згідно (Main Department of Statistics in the Khmelnytskyi region of Ukraine, n.d.)

З рис. 1 можна побачити що за останні 15 років поголів'я великої рогатої худоби і корів має загальну тенденцію зменшуватися. На графіку поголів'я великої рогатої худоби чітко видно тренд і сезонність.

В таблиці 1 представлені найкращі моделі SARIMAX для прогнозування поголів'я великої рогатої худоби і корів, використовуючи критерій AIC.

Таблиця 1 – Параметри моделі SARIMAX і значення AIC–критерію для часових рядів, що характеризують поголів'я великої рогатої худоби і корів.

	Model	AIC
Cattle	SARIMAX(0, 1, 0)x(0,1, 1, 12)	-4071.359
Cows	SARIMAX(1, 1, 9)x(1,1, 1, 12)	465.673

Джерело: розроблено автором

Термін прогнозування(Rolling window size comparison RWS) вибираємо 24 місяця, оскільки при такому значенні RWS маємо найменші оцінки RSME та MAPE.

Прогнозні щомісячні значення поголів'я великої рогатої худоби, тис. од.												
Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2023	211	217	234	248	240	227	220	221	236	245	236	218
2024	212	217	235	249	241	227	220	222	237	246	236	218
Прогнозні щомісячні значення поголів'я корів, тис. од.												
2023	121	122	122	122	122	122	122	121	121	120	120	120
2024	120	120	120	120	121	120	120	119	119	119	118	119

Джерело: розроблено автором

Прогнозування майбутнього аграрного сектору економіки, який є критично важливим компонентом, має важливе значення як для розвинутих країн, так і країн, що розвиваються. Крім того, можна буде визначити майбутню аграрну політику, спланувати інвестиції та вжити необхідних заходів. Найуспішніший прогноз, використовуючи щомісячні дані,

можна отримати шляхом удосконалення методів сезонного прогнозування. У цьому дослідженні підхід, який називається методом прогнозування SARIMAX, рекомендовано для прогнозування щомісячного поголів'я великої рогатої худоби і корів. Запропонований метод зосереджений на пошуку найбільш релевантних значень минулих спостережень шляхом ідентичних оцінок. Крім того, він виконує тести для всіх застосовних значень сезонного коефіцієнта. Таким чином ефективніше моделюється сезонність у наборі даних.

Література:

1. Atalan A. (2022) Forecasting drinking milk price based on economic, social, and environmental factors using machine learning algorithms. *Agribusiness*, 39(1), pp.214-241. DOI: 10.1002/agr.21773
2. Ediger V., Akar S. (2007) ARIMA forecasting of primary energy demand by fuel in Turkey. *Energy Policy*, Vol. 35, pp.1701–1708. DOI: 10.1016/j.enpol.2006.05.009
3. Main Department of Statistics in Khmelnytskyi Region. (n.d.). Retrieved from: <https://www.km.ukrstat.gov.ua/ukr/index.htm>