

## **СИСТЕМНИЙ ПІДХІД ДО АНАЛІЗУ ІНЦИДЕНТІВ В АВІАЦІЇ: ARIMA-МОДЕЛЬ ДЛЯ ДОВГОСТРОКОВИХ СПОСТЕРЕЖЕНЬ**

Авіаційна галузь є критично важливою для глобальної мобільності, однак її безпека залишається основним викликом через складність технічних систем і потенційні ризики інцидентів. Використання ARIMA-моделей у дослідженнях дозволяє проводити аналіз довгострокових трендів, виявляючи залежності даних про технічні несправності за минулі роки. Це допомагає розробляти заходи для зниження ризиків та запобігання аваріям. Завдяки здатності обробляти складні часові ряди, ARIMA інтегрується у системний підхід, який враховує взаємодію технічних, людських і екологічних факторів.

Мета даної роботи полягає у проведенні системного аналізу інцидентів в авіації за допомогою ARIMA-моделі, використовуючи багаторічні дані для прогнозування та мінімізації технічних ризиків, що впливають на безпеку польотів.

Аналіз часових рядів є одним з ключових методів для роботи з даними, що змінюються в часі, особливо у сферах, де потрібне прогнозування тенденцій або оцінка ризиків. Застосування моделей, таких як ARIMA, є класичним підходом, який дозволяє ефективно працювати зі стаціонарними часовими рядами завдяки поєднанню авторегресивних компонентів (AR), ковзного середнього (MA) і інтегрування для корекції нестационарності.

У сучасній науковій літературі розглядаються як традиційні підходи до аналізу часових рядів, так і інноваційні комбінації моделей ARIMA з методами машинного навчання. Наприклад, у роботі [1] підкреслено, що ARIMA добре справляється із сезонними змінами в даних, а також ефективно моделює короткострокові тенденції. Робота [2] описує порівняння цієї моделі з нейронними мережами, які краще враховують нелінійні взаємозв'язки, але потребують більше даних і складніші у налаштуванні. Особливий інтерес викликає поєднання ARIMA з іншими методами, наприклад, SARIMA для сезонних даних [3], чи гібридні підходи, що інтегрують ARIMA з LSTM-мережами для довгострокового прогнозування, про що йдеться в [4]. Це дозволяє покращити точність прогнозів у галузях, де дані мають складну структуру. Зокрема, у сфері авіації ARIMA може використовуватися для аналізу технічних інцидентів, прогнозування частоти несправностей, а також для виявлення аномалій у часових рядах, що описують параметри польотів. У

[5] показано, як часові ряди з даними про інциденти допомагають розробляти стратегії для запобігання аваріям.

Проведений аналіз робіт вказує на наявність невирішених проблем, пов'язаних з ризиками в авіації, що виникають через технічні неполадки. Однією з таких є прогнозування рідкісних подій, оскільки серйозні несправності трапляються нечасто, і тому важко побудувати точні прогнози для таких випадків. Крім того, нестационарність даних, тобто наявність трендів, є поширеною проблемою, оскільки більшість моделей вимагають, щоб дані були стабільними з часом. Існує також проблема обробки пропусків у даних, оскільки реальні набори часто містять відсутні значення. Використання моделі ARIMA дозволяє вирішити ці проблеми, оскільки вона добре підходить для прогнозування нестационарних часових рядів за допомогою попереднього диференціювання. Для рідкісних подій ARIMA здатна виявляти приховані закономірності, навіть коли ці події виникають нечасто. Крім того, ARIMA дозволяє працювати з пропущеними даними, застосовуючи відповідні техніки для їх інтерполяції. Тому ця модель може забезпечити точніші прогнози технічних несправностей в авіації.

Нами використана статистика інцидентів із залученням літаків Boeing 737 (усіх модифікацій) з 2009 по 2024 роки [6]. Дані базуються на доступній інформації та відображають загальну кількість інцидентів, а також ті випадки, причиною яких стали технічні несправності. Для оцінки трендів та виділення основних факторів ризику задіяно модель ARIMA. Розрахунки виконано за допомогою розробленого та реалізованого в MATLAB алгоритму, який включає наведені нижче етапи.

Етап 1. Завантаження даних.

Етап 2. Графічний аналіз. На цьому етапі створюються два графіки: перший показує зміни загальної кількості інцидентів по роках, другий – кількість технічних інцидентів по роках. Це дозволяє візуально оцінити тренди та коливання в даних.

Етап 3. Перевірка на стаціонарність. Для кожного з рядів (загальних і технічних інцидентів) виконується перевірка на стаціонарність за допомогою тесту Діккі-Фуллера. Якщо р-значення більше 0.05, то ряд не є стаціонарним (має тенденцію до зміни середнього чи варіації з часом).

Етап 4. Диференціація. Якщо дані не стаціонарні, то застосовується диференціація (різниця між поточним і попереднім значенням). Диференціація використовується для усунення тренду і забезпечення стаціонарності ряду.

Етап 5. Аналіз автокореляції. Цей процес дозволяє оцінити залежність між значеннями в часі, тобто, чи є кореляція між поточними і попередніми значеннями.

Етап 6. Побудова та оцінка моделей ARIMA. Створюються моделі ARIMA (0,1,1) для кожного з рядів.

ARIMA(0,1,1) означає модель без авторегресії, з одиничним диференціюванням і одним коефіцієнтом для ковзного середнього. Моделі оцінюються за допомогою методу максимальної ймовірності, і на основі цих моделей здійснюється прогноз на 5 років вперед для кожного ряду (загальних і технічних інцидентів).

Етап 7. Візуалізація прогнозу. Прогнозовані значення накладаються на дані і виводяться на графіках для кожного з рядів. Результати роботи програми наведені на рис. 1-3 і таблицях 1, 2.

На основі графічного аналізу (рис.1) робимо висновки: загальна кількість інцидентів за роками має тенденцію до зниження, з піками у 2012 і 2016 роках, інциденти через технічні несправності також зменшуються, з найбільшими показниками у 2012 і 2016 роках. Рис. 2 демонструє залежність між значеннями інцидентів у різні роки. Є видимі піки на різних лагах, що вказує на певну кореляцію між даними в різні роки. Це означає, що поточні показники інцидентів можуть бути пов'язані з попередніми роками, що вимагає моделювання за допомогою часових рядів для точних прогнозів. Також виявлено стабільність або повторюваність проблем з технікою в часі. Автокореляція для обох серій показує наявність деякої залежності, що підтверджує потребу у використанні моделей часових рядів для прогнозування.



Рисунок 1 – Гістограми частот кількості інцидентів

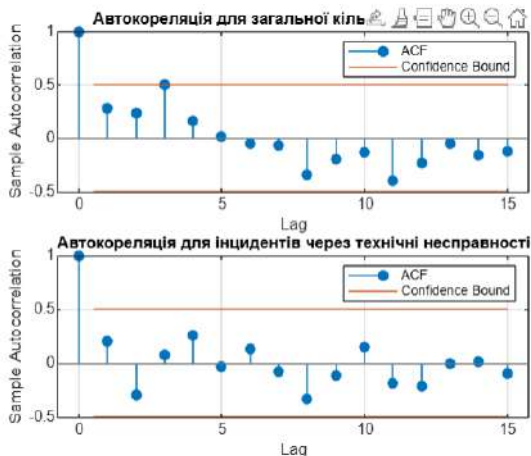


Рисунок 2 – Автокореляція для кількості інцидентів

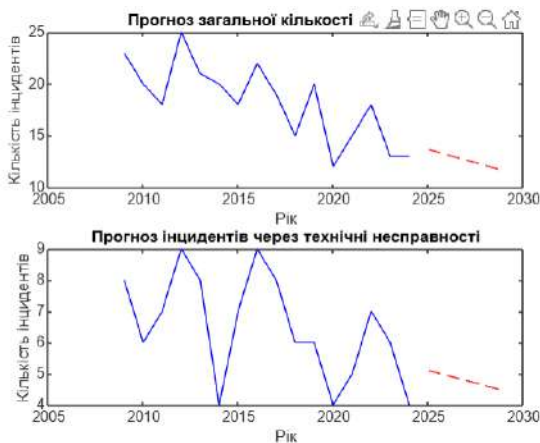


Рисунок 3 – Прогноз кількості інцидентів.

На рис. 3 продемонстровано прогноз кількості загальних та технічних інцидентів на наступні 5 років (2025-2029). Червона пунктирна лінія показує прогнози, які залишаються стабільними на рівні 14.33 загальних і 5.58 технічних інцидентів на рік. Це підтверджує, що модель ARIMA передбачає стабільну кількість інцидентів, без значних змін у майбутньому.

Тест Діккі-Фуллера для загальної кількості інцидентів має р-значення 0.96, що вказує на відсутність стаціонарності. Тест для технічних інцидентів має р-значення 0.15, що також свідчить про відсутність стаціонарності. Це

означає, що дані не є стаціонарними, тому застосовано диференціювання. Після диференціювання дані стали більш стабільними для подальшого аналізу, однак, через відсутність істотних змін у тренді, було прийнято використовувати модель ARIMA.

Моделі ARIMA(0,1,1) для обох серій були оцінені (табл.1,2), і прогнози для наступних 5 років показують наступні результати: прогноз для загальної кількості інцидентів: 14.33 (для кожного року прогнози однакові), прогноз для інцидентів через технічні несправності: 5.58 (також постійний прогноз).

Таблиця 1 – ARIMA(0,1,1) Model (Gaussian Distribution)

	Value	StandardError	TStatistic	PValue
<b>Constant</b>	-0.5185	0.12553	-4.1305	3.6194e-05
<b>MA{1}</b>	-1	0.29097	-3.4368	0.00058864
<b>Variance</b>	6.0244	4.1263	1.46	0.1443

Таблиця 2 – ARIMA(0,1,1) Model (Gaussian Distribution)

	Value	StandardError	TStatistic	PValue
<b>Constant</b>	-0.16087	0.03973	-4.0492	5.1399e-05
<b>MA{1}</b>	-1	0.34835	-2.8707	0.0040961
<b>Variance</b>	1.9893	1.1004	1.8078	0.070643

Загальна кількість інцидентів в майбутньому, згідно з прогнозом ARIMA, залишатиметься стабільною на рівні 14-15 інцидентів на рік. Інциденти через технічні несправності також, ймовірно, залишатимуться на рівні близько 5-6 інцидентів на рік, без значних змін. Відсутність значних змін у трендах вказує на стабільність ситуації в області технічних несправностей. Подальші кроки повинні включати моніторинг і дослідження можливих факторів, що можуть впливати на ці стабільні показники, наприклад, технічні інновації або зміни у стандартах безпеки.

### СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Kontopoulou, V.I., Panagopoulos, A.D., Kakkos, I., & Matsopoulos, G.K. "A Review of ARIMA vs. Machine Learning Approaches for Time Series Forecasting in Data Driven Networks." Future Internet, 2023. doi:10.3390/fi15080255.
2. Hyndman, R.J., & Athanasopoulos, G. "Forecasting: Principles and Practice." OTexts, 2021.
3. Adhikari, R., & Agrawal, R.K. "An Introductory Study on Time Series Modeling and Forecasting." LAP Lambert Academic Publishing, 2013. doi:10.48550/arXiv.1302.6613.

4. Box, G.E.P., Jenkins, G.M., & Reinsel, G.C. "Time Series Analysis: Forecasting and Control." Prentice Hall, 2015.
5. Chatfield, C. "The Analysis of Time Series: An Introduction." CRC Press, 2016.
6. Aviation Safety Network:<https://asn.flightsafety.org/>