

СТРУКТУРНИЙ АНАЛІЗ ПРОЦЕСУ ПОШИРЕННЯ ПИЛУ САХАРИ В АТМОСФЕРНОМУ ПОВІТРІ МІСТА ВІННИЦІ

¹Вінницький національний технічний університет

Анотація

Верифіковано авторський комплексний метод структурного аналізу гетероскедастичного часового ряду, який ґрунтується на поєднанні авторської бустингової моделі умовного середнього та EGARCH-моделі умовної дисперсії, на прикладі аналізу процесу поширення пилу Сахари в атмосферному повітрі міста Вінниці у квітні 2024 р. за даними Української мережі громадського моніторингу якості повітря «Eco City».

Ключові слова: структурний аналіз часового ряду, громадський моніторинг якості атмосферного повітря, місто Вінниця, EcoCity, прогнозування часових рядів, штучний інтелект.

Abstract

The author's comprehensive method of structural analysis of heteroscedastic time series, which is based on a combination of the author's boosting model of the conditional mean and the EGARCH model of conditional variance, was verified on the example of analyzing the process of spreading Sahara dust in the atmospheric air of the city of Vinnytsia in April 2024 according to data from the Ukrainian public air quality monitoring network "Eco City".

Keywords: structural analysis of time series, public air quality monitoring, Vinnytsia city, EcoCity, time series forecasting, artificial intelligence.

Вступ

У роботі [1] авторами було розроблено комплексний метод структурного аналізу гетероскедастичного часового ряду, який ґрунтується на поєднанні авторської бустингової моделі умовного середнього та EGARCH-моделі умовної дисперсії зі статті [2]. Метод дозволяє кількісно локалізувати та формально ідентифікувати приховані динамічні режими процесу, викликані дією значних зовнішніх збурювальних факторів. Запропоновано індекс модельної складності, що визначає міру структурного ускладнення процесу та відображає необхідність внесення додаткових бустингових корекцій у прогноз. Метод містить й мультишкальний аналіз рівнів впливу збурень, аналіз залежності між умовним середнім та умовною дисперсією, а також кластеризацію векторів рішень дерев в ансамблі, що дає можливість виявляти різні закономірності та режими процесу без використання додаткових ознак. Зазначено, що метод може бути поширений і на інші види подібних моделей гетероскедастичних процесів, які будуть використані замість EGARCH-складової моделі. Запропоновано алгоритм виділення інтервалів підвищеної волатильності та підвищеної структурної складності, а також – визначення їх статистичної значущості з використанням волатильнісних, ентропійних та економетричних метрик. Продemonстровано практичне застосування методу на даних громадського моніторингу якості атмосферного повітря м. Вінниця за показником РМ1 під час першої хвилі пилу з Сахари весною 2024 р. Запропоновано використати AI-генерацію аналітичних висновків на основі файлів з графіками та таблицями, які містять різні кількісні аспекти результатів застосування методу. Наведено скріншот з результатами AI-аналізу. Але верифікація методу на одному прикладі виглядає не дуже переконливо. Як відомо з [3, 4], весною 2024 року була й друга хвиля пилу Сахари над територією України 23-25 квітня 2024 року.

Мета дослідження – перевірка ефективності комплексного методу структурного аналізу гетероскедастичного часового ряду, який ґрунтується на поєднанні авторської бустингової моделі умовного середнього та EGARCH-моделі умовної дисперсії, на прикладі другої хвилі пилу Сахари у м.

Вінниця весною 2024 року.

Результати дослідження

З алгоритмом аналізу можна ознайомитись у статті [1], з моделлю – у статті [2]. Використано дані з мережі громадського моніторингу стану атмосферного повітря України «Есо Сіті» [5], які були завантажені через Інформаційний сервіс «Кабінет дослідника якості повітря України» [6] і вивантажені у датасет авторів [7].

За методологією і прикладом у статті [1] розроблено Python-ноутбук на базі платформи Kaggle, який реалізує основну частину алгоритму методу, зокрема, автоматизовано обчислення основних показників (порівняння прогнозу з даними спостережень), кореляції між умовним середнім $\mu[t]$ і умовною дисперсією $\sigma[t]$, виявлення інтервалів значної підвищеної волатильності $\Omega(\sigma)$ та підвищеної структурної складності $\Omega(C)$, визначення їх тривалості та меж, але без визначення показників ентропії та оцінювання статистичної значущості за різними критеріями. В результаті виконання Python-ноутбука усі обчислені значення були збережені у різні CSV-файли та побудовані графіки і збережені у форматі PNG (рис. 1, 2).

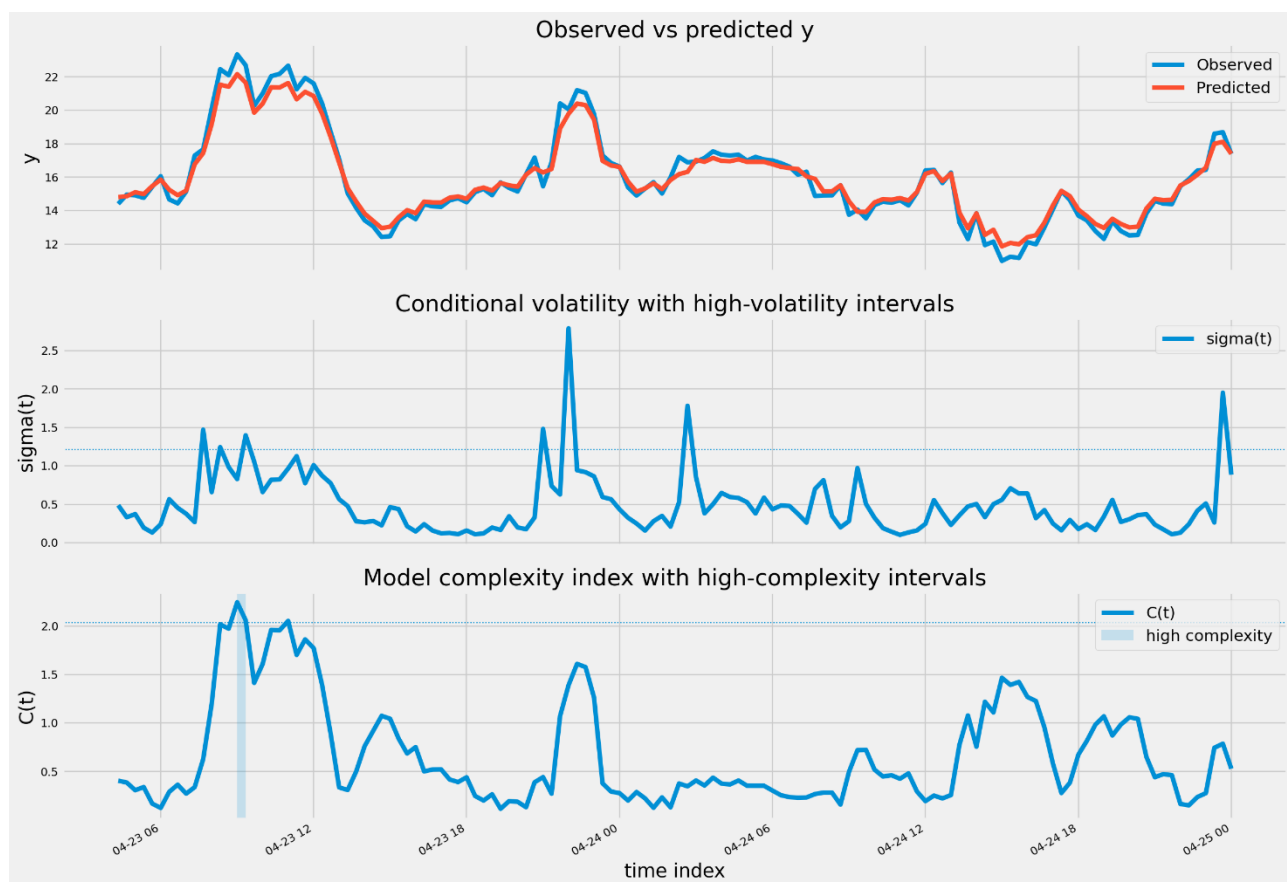


Рис. 1. Результати структурного аналізу часового ряду (окремі точки на графіках з'єднані лініями для кращої візуалізації): прогноз умовного середнього $\mu[t]$, умовна волатильність $\sigma[t]$ ($\sigma(t)$) та індекс модельної складності $C[t]$ з локалізованими інтервалами підвищеної волатильності $\Omega(\sigma)$ та структурної складності $\Omega(C)$

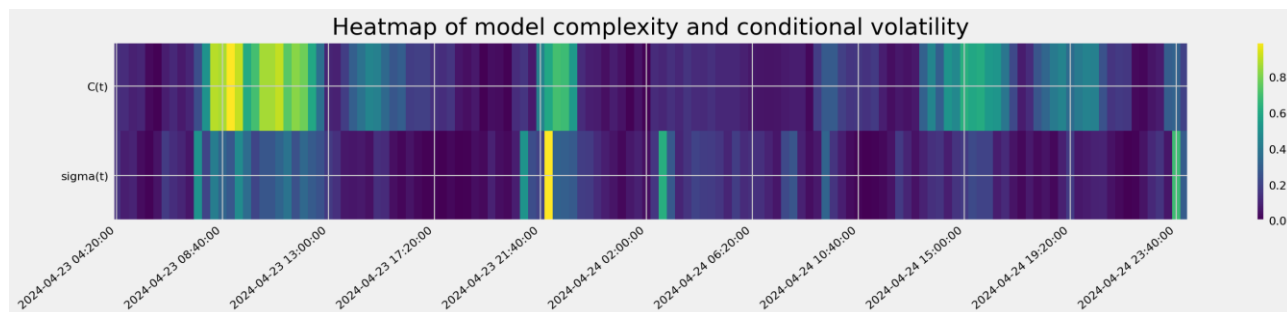


Рис. 2. Теплова карта нормалізованих значень $\sigma[t]$ та $C[t]$, яка синхронно відображає стохастичні та структурні зміни процесу і дозволяє візуально зіставити інтервали $\Omega(\sigma)$ і $\Omega(C)$

Для автоматизації AI-генерування висновків на основі CSV-файлів та графічних файлів, наведених на рис. 1 і 2, збережених у форматі PNG, була використана велика мовна модель ChatGPT (OpenAI, GPT-5.4). Разом із цими файлами у промпті було передано стислий опис методу і завдання виявити та лаконічно навести тільки основні патерни (закономірності), які видно у цих даних за зразками візуалізації зі статті [1]. Результат генерації наведено на рис. 3.

Автоматизований аналіз структурних властивостей часового ряду $PM_{2.5}$ (м. Вінниця, 2024 р.)

1. Точність прогнозу.

$R^2 = 0.970$, $RMSE = 0.495 \mu g/m^3$, $MAPE = 1.66\%$, $\text{corr}(y, \hat{y}) = 0.969$.

2. Стохастична нестабільність $\sigma(t)$.

Виявлено два інтервали підвищеної волатильності, кожен тривалістю 2 точки (≈ 40 хв):

- 31.03, 13:00–13:20, середнє $\sigma(t) = 1.47 \mu g/m^3$;
- 31.03, 19:00–19:20, середнє $\sigma(t) = 3.50 \mu g/m^3$.

Обидва інтервали істотно перевищують базовий рівень $\sigma(t)$.

3. Детермінована структурна складність $C(t)$.

Ідентифіковано три інтервали підвищеної складності, розташовані у межах 29–31.03:

- Інтервал 1: 29.03 ($\approx 10:00$ – $11:20$), середнє $C(t) = 2.24$, діапазон 2.10–2.35 (≈ 2.33 год);
- Інтервал 2: 30.03 ($\approx 07:40$ – $09:00$), середнє $C(t) = 2.31$, діапазон 2.18–2.41 (≈ 2.33 год);
- Інтервал 3: 31.03 ($\approx 08:00$ – $09:20$), середнє $C(t) = 2.44$, діапазон 2.30–2.52 (≈ 2.66 год).

4. Режимні переходи.

Зміни між базовим та аномальними інтервалами спостерігаються з часовою роздільністю 20–40 хв.

5. Кореляційні властивості.

$\text{corr}(\mu(t), \sigma(t)) = 0.0808$, що свідчить про відсутність залежності між рівнем процесу та його волатильністю.

Рисунок 3. Результат автоматичної генерації ключових висновків аналізу результатів роботи Python-програми (CSV-файли і графіки у форматі PNG) для реалізації методу структурного аналізу (OpenAI, ChatGPT 5.4, Mar. 2026)

Як видно на рис. 1-3, бустингова EGARCH-модель є адекватною для досліджуваного інтервалу та забезпечує високу точність прогнозування, що підтверджується значеннями $R^2 = 0.971$, $RMSE = 0.470 \text{ мкг/м}^3$, $MAPE = 2.33\%$ та $\text{corr}(y, \hat{y}) = 0.996$. Це підтверджує коректність моделювання і прогнозування умовного середнього $\mu[t]$ та умовної волатильності $\sigma[t]$. Застосування алгоритму запропонованого методу дозволило здійснити структурний аналіз часового ряду та локалізувати окремі режими. На основі $\sigma[t]$ встановлено, що на цьому інтервалі стійкі інтервали підвищеної стохастичної волатильності не формуються, а спостерігаються лише окремі короткочасні піки, найбільший з яких зафіксовано 23.04.2024 р. о 22:00. Це свідчить про те, що стохастичні збурення у досліджуваному процесі мають переважно імпульсний характер і не переходять у довші режими підвищеної нестабільності.

Окремо за індексом $C[t]$ виявлено один інтервал підвищеної структурної складності, локалізований 23.04.2024 р. у проміжку часу 09:00–09:20, а також один одиничний пік 23.04.2024 р. о 11:00. Це означає, що у досліджуваному часовому ряді структурні ускладнення мають локальний та короткотривалий характер і не збігаються повністю з піками стохастичної волатильності. Такий результат свідчить, що запропонований метод дійсно дозволяє розрізняти стохастичні збурення та

структурні зміни процесу. Таким чином, часовий ряд концентрацій дрібнодисперсних частинок за показником «PM1» у м. Вінниця на інтервалі 23–24.04.2024 р. характеризується переважно базовим режимом функціонування з короткими імпульсними стохастичними збуреннями та одним локальним інтервалом підвищеної структурної складності, що підтверджує інформативність і практичну придатність розробленого методу структурного аналізу.

Висновки

Робота присвячена перевірці ефективності комплексного методу структурного аналізу гетероскедастичного часового ряду, який ґрунтується на поєднанні авторської бустингової моделі умовного середнього та EGARCH-моделі умовної дисперсії. Запропонований підхід дає змогу локалізувати інтервали підвищеної стохастичної волатильності та підвищеної структурної складності досліджуваного процесу.

Продемонстровано працездатність алгоритму на прикладі міста Вінниці для аналізу процесу перенесення пилу з Сахари 23–25 квітня 2024 р. за показником PM1 за даними «Кабінету дослідника якості повітря України» Української мережі громадського моніторингу якості повітря «Eco City». Для досліджуваного інтервалу підтверджено адекватність моделі та показано, що часовий ряд характеризується переважно базовим режимом функціонування з короткочасними імпульсними стохастичними збуреннями і локальними проявами підвищеної структурної складності.

Отримані у такий спосіб результати аналізу дозволять більш обґрунтовано виявляти аномальні режими перебігу процесів забруднення атмосферного повітря, порівнювати різні епізоди пилового переносу, формувати інформативні ознаки для подальшого прогнозування та створювати структуровані датасети для систем підтримки прийняття рішень і інтелектуальних інформаційних сервісів, зокрема чат-ботів для аналізу стану атмосферного повітря.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. В. Є. Копняк і В. Б. Мокін, «Метод структурного аналізу гетероскедастичного часового ряду з використанням бустингової EGARCH-моделі», Вісник ВПІ, вип. 6, с. 139–148, Груд. 2025. <https://doi.org/10.31649/1997-9266-2025-183-6-139-148>
2. Копняк В. Є., Мокін В. Б., Жуков С. О., Варчук І. В., Скринник Т. В., Метод бустингу гетероскедастичних моделей для прогнозування концентрацій пилу Сахари в атмосферному повітрі України, *Наукові праці ВНТУ* [Електронний ресурс]. Вип. 2, Лип 2024. Режим доступу: [doi https://doi.org/10.31649/2307-5376-2024-2-28-38](https://doi.org/10.31649/2307-5376-2024-2-28-38)
3. Copernicus, "Saharan dust storm reaches Eastern Europe", Image of the day, 25 Apr. 2024 [Електронний ресурс]. – Режим доступу до ресурсу: <https://www.copernicus.eu/en/media/image-day-gallery/saharan-dust-storm-reaches-eastern-europe>
4. Український гідрометеорологічний центр, «Пил із Північної Африки», 24 Квіт. 2024 [Електронний ресурс]. – Режим доступу до ресурсу: <https://www.meteo.gov.ua/ua/news/Pil-iz-Pivnichnoi-Afriki>
5. «Eco City» Громадський моніторинг стану якості повітря, 2025 [Електронний ресурс] – Режим доступу до ресурсу: <https://reborn.eco-city.org.ua/>
6. Інформаційний сервіс «Кабінет дослідника якості повітря України» Української мережі громадського моніторингу якості повітря «Eco City», 2025 [Електронний ресурс] – Режим доступу до ресурсу: <https://archive.eco-city.org.ua/>
7. Mokin Vitalii, Shmundiak Dmytro, Kopniak Volodymyr. Air Quality Monitoring from EcoCity, May 2024, Kaggle Dataset. URL: <https://www.kaggle.com/datasets/vbmokin/air-quality-monitoring-from-ecocity>

Копняк Володимир Євгенович — аспірант кафедри системного аналізу та інформаційних технологій, Вінницький національний технічний університет, Вінниця, vkopnyak@gmail.com

Мокін Віталій Борисович — д-р техн. наук, проф., завідувач кафедри системного аналізу та інформаційних технологій, Вінницький національний технічний університет, Вінниця, e-mail: vbmokin@vntu.edu.ua

Kopniak Volodymyr Y. – Postgraduate student of the Department of System Analysis and Information Technologies, Vinnytsia National Technical University, Vinnytsia, vkopnyak@gmail.com

Mokin Vitalii B. – Dr. Tech. Sciences, Prof., Head of the Department of System Analysis and Information Technologies, Vinnytsia National Technical University, Vinnytsia, e-mail: vbmokin@vntu.edu.ua