

ПОПЕРЕДНЯ ОБРОБКА ТА РОЗВІДУВАЛЬНИЙ АНАЛІЗ ДАНИХ МОНІТОРИНГУ АТМОСФЕРНОГО ПОВІТРЯ МІСТА ВІННИЦІ

¹ Вінницький національний технічний університет

Анотація

Проведено розвідувальний аналіз даних моніторингу атмосферного повітря міста Вінниці протягом одного року. Дослідження охоплює концентрації завислих речовин, діоксиду сірки, оксиду вуглецю, діоксиду азоту, фтористого водню, аміаку та формальдегіду. Виконано аналіз описових статистик, форми розподілів, викидів, кореляційних залежностей, та автокореляційної структури часових рядів. Встановлено наявність суттєвої асиметрії та екстремальних значень для більшості забруднювачів, що потребує застосування процедур попередньої обробки даних. Аналіз рангових кореляцій показав відсутність сильних взаємозв'язків між концентраціями досліджуваних речовин, а автокореляційний аналіз виявив різний ступінь часової залежності окремих забруднювачів.

Ключові слова: атмосферне повітря, моніторинг довкілля, розвідувальний аналіз даних, часові ряди, кореляційний аналіз, автокореляція, забруднення повітря.

Abstract

An exploratory data analysis of atmospheric air quality monitoring data collected in Vinnytsia during one year was conducted. The study considers concentrations of total suspended particles, sulfur dioxide, carbon monoxide, nitrogen dioxide, hydrogen fluoride, ammonia, and formaldehyde. Descriptive statistics, distribution characteristics, outlier analysis, correlation structure, and temporal dependencies were investigated. Significant right-skewness and extreme observations were identified for most pollutants, indicating the necessity of outlier treatment procedures. Spearman correlation analysis revealed weak relationships between pollutant concentrations, while autocorrelation analysis demonstrated different levels of temporal persistence among pollutants.

Keywords: air quality monitoring, exploratory data analysis, environmental data, time series, autocorrelation, correlation analysis, air pollution.

Вступ

Моніторинг атмосферного повітря є одним із інструментів оцінювання екологічного стану урбанізованих територій [1]. Концентрації забруднювальних речовин формуються під впливом великої кількості факторів, серед яких антропогенні викиди, транспортні потоки, промислові процеси та метеорологічні умови. Для забезпечення адекватного моделювання та прогнозування якості повітря необхідним є попередній розвідувальний аналіз даних, який дозволяє виявити структуру даних, аномалії, закономірності та потенційні проблеми набору даних.

Метою роботи є проведення комплексного розвідувального аналізу показників забруднення атмосферного повітря міста Вінниці та виявлення статистичних особливостей часових рядів концентрацій основних забруднювальних речовин [2].

Підготовка та первинний аналіз даних

У дослідженні використано результати вимірювань концентрацій семи забруднювальних речовин за період з 21 грудня 2024 року по 20 грудня 2025 року [3]. Найбільша середня концентрація спостерігалась для оксиду вуглецю, тоді як найменші значення характерні для діоксиду сірки та фтористого водню. Аналіз стандартних відхилень показав значну мінливість концентрацій завислих речовин, формальдегіду, аміаку та діоксиду азоту. Для оцінювання форми розподілів було розраховано коефіцієнти асиметрії та ексцесу. Встановлено, що більшість досліджуваних показників мають суттєве відхилення від нормального розподілу. Основні характеристики результатів спостережень наведені у таблиці 1.

Таблиця 1 - Основні характеристики набору даних

	ЗР	SO₂	CO	NO₂	HF	NH₃	CH₂O
Кількість	197	198	193	196	148	196	166
E	0.1	0.0007	0.91	0.089	0.007	0.019	0.24
σ	0.35	0.0099	0.44	0.16	0.013	0.093	0.42
γ	11.8	13.9	0.47	7.62	10.3	12.7	2.28
κ	153.5	192.5	2.78	63.7	116.6	169.5	3.8
max	4.8	0.14	3	1.68	0.165	1.28	0.18
P ₉₉	0.45	0.0001	1.84	0.71	0.016	0.148	0.16

Особливо виражена правостороння асиметрія спостерігається для SO₂, NH₃, завислих речовин та HF. Аналогічно високі значення ексцесу для цих показників свідчать про наявність рідкісних, але екстремально високих концентрацій. Отримані результати можуть бути пов'язані як із короточасними епізодами підвищеного забруднення атмосферного повітря, так і з наявністю аномальних спостережень у даних моніторингу. Для завислих речовин, діоксиду азоту, фтористого водню та аміаку максимальні значення значно перевищували відповідні значення 99-го перцентилля. У зв'язку з цим було застосовано процедуру вінзорування з верхньою межею на рівні P₉₉.

Після обробки викидів статистичні характеристики досліджуваних рядів суттєво змінилися. Найбільший ефект спостерігався для завислих речовин та фтористого водню. Для HF коефіцієнт асиметрії зменшився до 0,5, а ексцес набув від'ємного значення, що свідчить про наближення розподілу до симетричного вигляду. Для завислих речовин коефіцієнт асиметрії зменшився до 1,59, а ексцес – до 1,89. Результати обробки наведені в таблиці 2.

Таблиця 2 - Основні характеристики змінених даних

	ЗР	NO₂	HF	NH₃
E	0.085	0.081	0.006	0.012
σ	0.11	0.097	0.004	0.02
γ	1.59	5.43	0.5	4.5
κ	1.89	30.7	-0.74	24.2

Разом з тим для NH₃ та NO₂ навіть після усунення екстремальних значень зберігалася значна асиметрія розподілу. Додаткове логарифмічне перетворення не призвело до суттєвого покращення показників нормальності. Це свідчить про те, що правостороння асиметрія є природною властивістю розподілів концентрацій цих забруднювачів.

Окремої уваги потребують результати аналізу SO₂. Значення 95-го перцентилля дорівнює нулю, що може свідчити або про фактичну відсутність забруднення діоксидом сірки, або про концентрації нижче межі виявлення вимірювального обладнання. За таких умов статистичні висновки щодо цього показника необхідно інтерпретувати з обережністю.

Результати дослідження

Для дослідження взаємозв'язків між концентраціями забруднювачів було використано ранговий коефіцієнт кореляції Спірмена. Отримані значення знаходяться в межах від -0,15 до 0,22, що відповідає дуже слабким кореляційним зв'язкам (рисунок 1). Відсутність значущих кореляцій може свідчити про відсутність єдиного домінуючого джерела забруднення атмосферного повітря. Різні речовини можуть формуватися внаслідок різних фізико-хімічних процесів та надходити з різних

джерел. Зокрема, CO та NO₂ переважно пов'язані з транспортними викидами, NH₃ може бути пов'язаний із сільськогосподарською діяльністю та біогенними процесами, HF – з локальними промисловими джерелами, а завислі речовини значною мірою залежать від дорожнього пилу, будівельної діяльності та погодних умов. Отримані результати також можуть свідчити про суттєвий вплив метеорологічних факторів, які по-різному впливають на концентрації окремих забруднювачів [4].

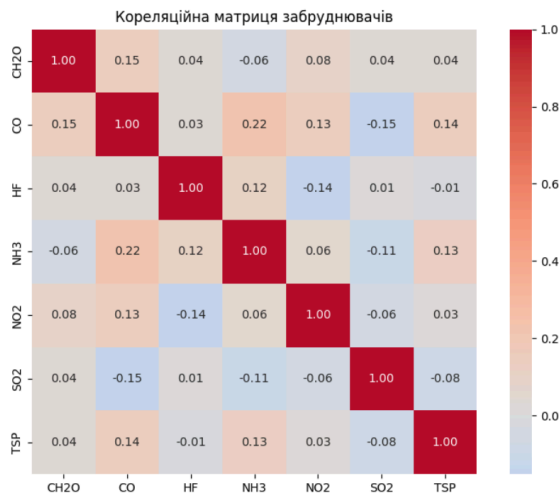


Рис. 1. Кореляційна матриця забруднювачів

Додатково проведено аналіз автокореляційної структури часових рядів. Найбільше значення автокореляції першого порядку було отримано для оксиду вуглецю, що свідчить про виражену часову інерційність процесу. Для аміаку та завислих речовин отримано помірні значення автокореляції. Для формальдегіду, фтористого водню, діоксиду азоту та діоксиду сірки автокореляція першого порядку була близькою до нуля, що свідчить про відсутність значущої часової пам'яті. Аналіз автокорелограм показав, що для CO статистично значущі кореляції зберігаються до 10-12 діб (рисунок 2), що вказує на наявність процесів накопичення забруднення та стабільних джерел викидів. Для NH₃ часові залежності мають короткостроковий характер, а для завислих речовин автокореляція швидко згасає.

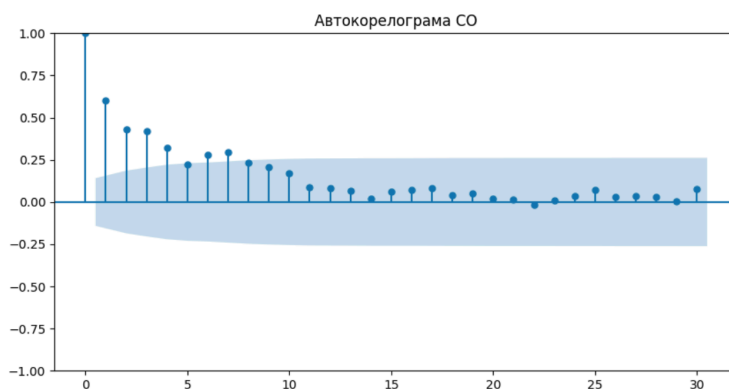


Рис. 2. Автокорелограма оксиду вуглецю

Висновки

Проведений розвідувальний аналіз даних моніторингу атмосферного повітря міста Вінниці дозволив виявити низку статистичних особливостей досліджуваних показників. Встановлено, що більшість концентрацій забруднювачів характеризуються вираженою правосторонньою асиметрією та наявністю екстремальних значень. Для зменшення впливу викидів доцільним виявилось застосування вінзорування на рівні 99-го перцентиля.

Кореляційний аналіз не виявив сильних взаємозв'язків між досліджуваними речовинами, що

свідчить про багатфакторний характер формування якості атмосферного повітря та відсутність єдиного домінуючого джерела забруднення. Автокореляційний аналіз показав різний ступінь часової залежності для окремих забруднювачів. Найбільш виражену інерційність продемонстрував оксид вуглецю, тоді як для більшості інших речовин часові залежності є слабкими або незначущими.

Після завершення збору даних протягом декількох років буде доцільно дослідити сезонні залежності величин. Отримані результати можуть бути використані як основа для подальшого моделювання, прогнозування концентрацій забруднювачів та оптимізації регіональних систем моніторингу атмосферного повітря.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Про охорону атмосферного повітря [Електронний ресурс] : Закон України від 16.10.1992 № 2707-ХІІ : станом на 8 серп. 2025 р. – Режим доступу: <https://zakon.rada.gov.ua/laws/show/2707-12#Text> (дата звернення: 30.05.2026).

2. Шмундяк Д. О. Системний аналіз стану атмосферного повітря регіону, з урахуванням впливу аномалій [Електронний ресурс] : монографія / Д. О. Шмундяк, В. Б. Мокін, Є. М. Крижановський. – Вінниця : ВНТУ, 2025. – 169 с. – Режим доступу: <https://press.vntu.edu.ua/index.php/vntu/catalog/book/925>.

3. Крижановський Є. М. Підхід до збору та візуалізації даних державного моніторингу стану атмосферного повітря міста Вінниці [Електронний ресурс] / Є. М. Крижановський, М. В. Швець // Матеріали LV Всеукраїнської науково-технічної конференції підрозділів ВНТУ (ВНТКП ВНТУ-2026), Вінниця, 24–27 берез. 2026 р. – [Б. м.], 2026. – Режим доступу: <https://conferences.vntu.edu.ua/index.php/all-fksa/all-fksa-2026/paper/view/27366>.

4. Sources and emissions of air pollutants in Europe [Електронний ресурс] // European Environment Agency (EEA). – Режим доступу: <https://www.eea.europa.eu/en/analysis/publications/air-quality-in-europe-2022/sources-and-emissions-of-air-pollutants-in-europe> (дата звернення: 30.05.2026).

Євгеній Миколайович Крижановський – к.т.н., доцент, доцент кафедри системного аналізу та інформаційних технологій, e-mail: kruzhan@gmail.com

Швець Михайло Владиславович – аспірант кафедри системного аналізу та інформаційних технологій, e-mail: misha.shvets02@gmail.com

Kryzhanovskiy Yevhenii M. – Cand. Sc. (Eng.), Associate Professor, Associate Professor of the Chair of System Analysis and Information Technologies, e-mail: kruzhan@gmail.com

Shvets Mykhailo V. – Post-Graduate Student, of the Chair of System Analysis and Information Technologies, e-mail: misha.shvets02@gmail.com