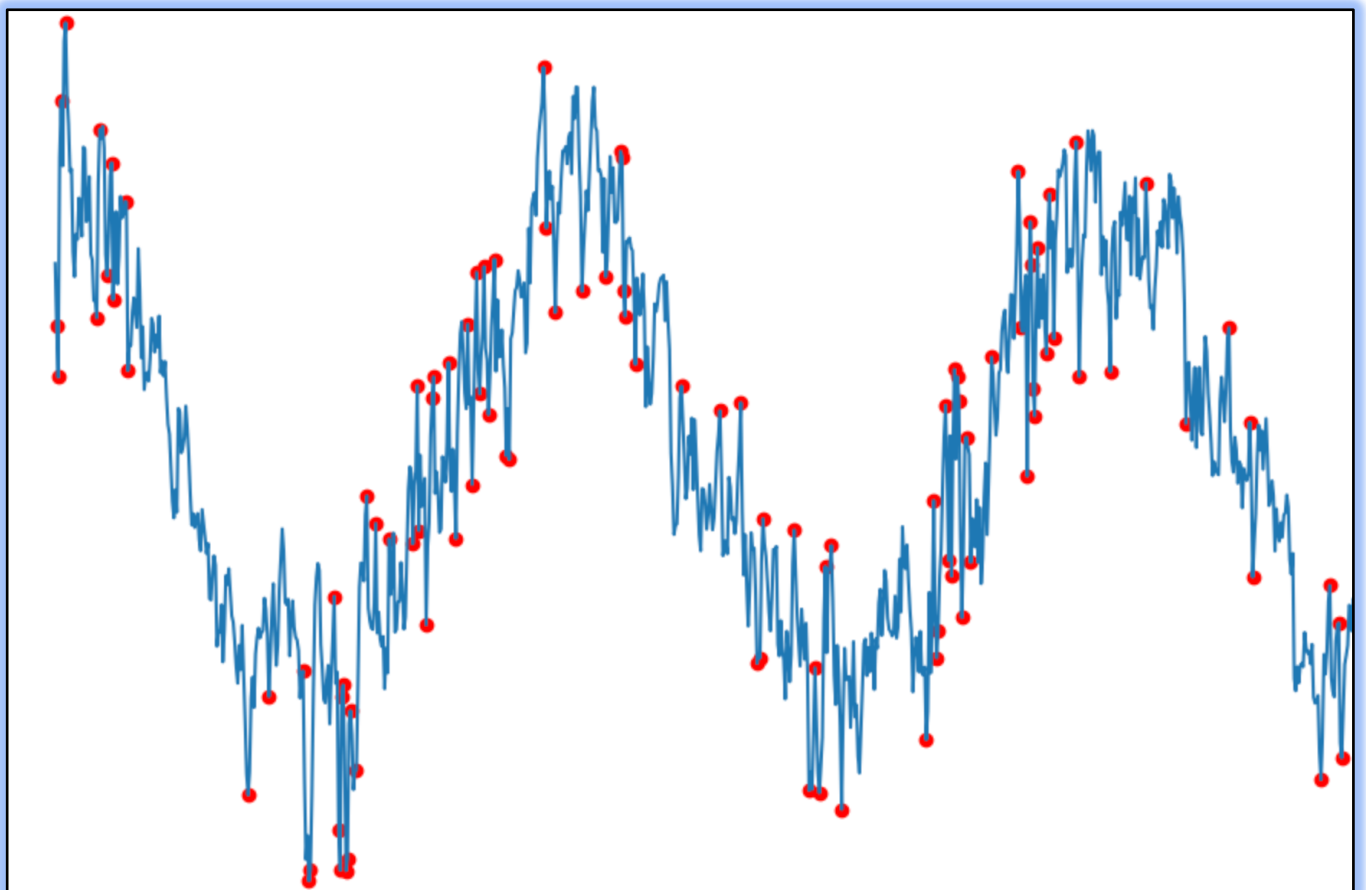


Д. О. ШМУНДЯК, В. Б. МОКІН, Є. М. КРИЖАНОВСЬКИЙ

**СИСТЕМНИЙ АНАЛІЗ СТАНУ
АТМОСФЕРНОГО ПОВІТРЯ РЕГІОНУ, З
УРАХУВАННЯМ ВПЛИВУ АНОМАЛІЙ**



Міністерство освіти і науки України
Вінницький національний технічний університет

Д. О. Шмундяк, В. Б. Мокін, Є. М. Крижановський

**СИСТЕМНИЙ АНАЛІЗ СТАНУ
АТМОСФЕРНОГО ПОВІТРЯ РЕГІОНУ,
З УРАХУВАННЯМ ВПЛИВУ
АНОМАЛІЙ**

Монографія

Вінниця
ВНТУ
2025

УДК 004.9+502/504

Ш 75

Рекомендовано до видання Вченою радою Вінницького національного технічного університету Міністерства освіти і науки України (протокол № 7 від 26.12.2024 р.).

Рецензенти:

Кветний Р. Н., член-кор. НАПН України, доктор технічних наук, професор, ВНТУ

Литвин В. В., доктор технічних наук, професор, НУ «Львівська політехніка»

Шмундяк, Д. О.

Ш 75 Системний аналіз стану атмосферного повітря регіону, з урахуванням впливу аномалій: монографія [Електронний ресурс] / Д. О. Шмундяк, В. Б. Мокін, Є. М. Крижановський. – Вінниця : ВНТУ, 2025. – (PDF, 169 с.)

ISBN 978-617-8163-54-9 (PDF)

В монографії представлена технологія системного аналізу стану атмосферного повітря регіону, з урахуванням аномалій, з використанням інтелектуальних моделей прогнозування даних спостережень та розкриття основних видів невизначеності, яка дозволяє підвищити точність прогнозування цього стану та підвищити ефективність системи управління станом атмосферного повітря в регіоні. Розрахована на студентів та аспірантів, науковців та практиків у сфері системного аналізу та екології.

УДК 004.9+502/504

ISBN 978-617-8163-54-9 (PDF)

© Д. Шмундяк, В. Мокін, Є. Крижановський, 2025

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ ТА СКОРОЧЕНЬ.....	6
ВСТУП	7
РОЗДІЛ 1 ОГЛЯД ТА АНАЛІЗ МЕТОДІВ І ЗАСОБІВ ПРОГНОЗУВАННЯ СТАНУ АТМОСФЕРНОГО ПОВІТРЯ	10
1.1. Стан атмосферного повітря в Україні та система моніторингу для його аналізу.....	10
1.1.1. Показники забруднення атмосферного повітря	10
1.1.2. Сучасний рівень забруднення атмосферного повітря в Україні	11
1.1.3. Державні механізми контролю та покращення стану атмосферного повітря.....	13
1.1.4. Проекти громадського моніторингу	14
1.1.5. Індекс якості повітря	18
1.2. Методологія системного аналізу процесів у складних системах.....	19
1.3. Характеристика об'єкта досліджень з позицій системного аналізу	24
1.3.1. Формалізація системи управління стану атмосферного повітря як складної системи	24
1.3.2. Декомпозиція складної системи за різними критеріями.....	25
1.3.3. Класифікація видів невизначеності	26
1.3.4. Підходи та методи щодо розкриття видів невизначеності.....	27
1.4. Огляд інтелектуальних моделей, методів і засобів прогнозування даних часових рядів	30
1.5. Огляд методів і засобів автоматизованого пошуку аномалій в часових рядах	33
1.5.1. Аналіз відомих методів пошуку аномалій	34
1.5.2. Огляд засобів на Python для автоматизації пошуку аномалій в часових рядах	39
1.6. Систематизація методів розкриття невизначеності для різних видів декомпозиції та постановка задач дослідження.....	42
1.7. Висновки до розділу 1.....	46
РОЗДІЛ 2 СИСТЕМНИЙ АНАЛІЗ СТАНУ АТМОСФЕРНОГО ПОВІТРЯ З УРАХУВАННЯМ ВПЛИВУ АНОМАЛІЙ	47
2.1. Математична постановка задачі системного аналізу	47

2.2. Метод пошуку можливих джерел підвищеного забруднення на основі аналізу аномалій	48
2.3. Метод формування системи критеріїв для оптимального розміщення станцій мережі моніторингу	56
2.3.1. Визначення екстремальних значень показників про стан атмосферного повітря.....	57
2.3.2. Визначення мінімальної кількості постів і складу показників спостережень за вітчизняним та європейським законодавством.....	58
2.3.3. Формування системи критеріїв для вибору параметрів місць постів спостережень за станом атмосферного повітря, з урахуванням екстремальних значень показників	62
2.3.4. Приклад застосування з використанням методу аналізу ієрархій.....	66
2.4. Висновки до розділу 2.....	69
РОЗДІЛ 3 ПРОГНОЗУВАННЯ СТАНУ АТМОСФЕРНОГО ПОВІТРЯ З УРАХУВАННЯМ ПІКОВИХ ЗНАЧЕНЬ	70
3.1. Метод ідентифікації параметрів сезонності часового ряду ...	70
3.1.1. Дослідження впливу параметрів моделі Prophet на точність прогнозування стану атмосферного повітря.....	70
3.1.2. Ідентифікація порядку ряду Фур'є для показника стану атмосферного повітря.....	73
3.1.3. Розроблення нового методу ідентифікації параметрів сезонності	75
3.2. Метод ідентифікації локальних аномалій часового ряду	82
3.3. Пошук взаємовпливів показників стану атмосферного повітря на основі байєсівського моделювання.....	93
3.4. Алгоритм системного аналізу стану атмосферного повітря регіону.....	98
3.5. Висновки до розділу 3.....	101
РОЗДІЛ 4 ОПТИМІЗАЦІЯ СИСТЕМИ МОНІТОРИНГУ ЗОНИ «ВІННИЦЬКА».....	102
4.1. Системний аналіз стану атмосферного повітря у зоні «Вінницька»	102
4.2. Оптимізація мережі спостережень та оцінювання методом системного аналізу	114
4.3. Висновки до розділу 4.....	122
ВИСНОВКИ.....	123
ПЕРЕЛІК ЛІТЕРАТУРИ.....	125

ДОДАТОК А ГРОМАДСЬКІ ПРОЄКТИ МОНИТОРИНГУ ЯКОСТІ АТМОСФЕРНОГО ПОВІТРЯ	130
ДОДАТОК Б ЄВРОПЕЙСЬКИЙ ІНДЕКС ЯКОСТІ ПОВІТРЯ	135
ДОДАТОК В УКРАЇНСЬКИЙ ІНДЕКС ЯКОСТІ ПОВІТРЯ.....	137
ДОДАТОК Г ФОРМАЛІЗАЦІЯ ІНФОРМАЦІЇ У ВІДПОВІДЯХ ЗАПИТІВ НА ПУБЛІЧНУ ІНФОРМАЦІЮ, ЗБИРАННЯ ЯКОЇ ВИМАГАЄ НАКАЗ МІНДОВКІЛЛЯ №147 ВІД 25.02.21.....	140
ДОДАТОК Д РОЗКРИТТЯ НЕВИЗНАЧЕНОСТІ	141
ДОДАТОК Е ЗАГАЛЬНА ХАРАКТЕРИСТИКА ЗОНИ «ВІННИЦЬКА».....	142
ДОДАТОК Є АНАЛІЗ ОСНОВНИХ ДЖЕРЕЛ ЗАБРУДНЕННЯ АТМОСФЕРНОГО ПОВІТРЯ ЗОНИ «ВІННИЦЬКА»	147
ДОДАТОК Ж АНАЛІЗ ЯКОСТІ АТМОСФЕРНОГО ПОВІТРЯ ТА ВИБІР РЕЖИМІВ СПОСТЕРЕЖЕНЬ.....	150
ДОДАТОК З ДИНАМІКА ВИКИДІВ ОСНОВНИХ ЗАБРУДНЮВАЛЬНИХ РЕЧОВИН ВІД СТАЦІОНАРНИХ ДЖЕРЕЛ НА ТЕРИТОРІЇ ЗОНИ «ВІННИЦЬКА».....	155
ДОДАТОК І ЗАПЛАНОВАНІ ЗАХОДИ ЩОДО МОДЕРНІЗАЦІЇ МЕРЕЖІ СПОСТЕРЕЖЕНЬ	163
ІНФОРМАЦІЯ ПРО АВТОРІВ.....	166

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ ТА СКОРОЧЕНЬ

AQI	– Air quality index
ARIMA	– Autoregressive Integrated Moving Average
Prophet	– модель для моделювання і прогнозування часових рядів
DAG	– Directed acyclic graph
EEA	– European Environment Agency
ВНТУ	– Вінницький національний технічний університет
ГДК	– Гранично допустима концентрація
ЄС	– Європейський союз
КМУ	– Кабінет Міністрів України
МОЗ	– Міністерство охорони здоров'я України

ВСТУП

Зміна клімату все більше стає відчутною і це проявляється в появі природних аномалій щодо температурного режиму атмосферного повітря, опадів, сили і напрямку вітру. Неочікувані циклони та антициклони спричиняють перенесення як вологи в атмосфері, так і забруднювальних речовин. У свою чергу, це спричиняє непередбачувані зміни стану атмосферного повітря. Крім того, в Україні в останні роки спостерігаються зміни у промисловому і сільськогосподарському секторі, пов'язані з переміщенням виробництв. А це спричиняє появу викидів антропогенного характеру, які не є характерними для певних місць у певний час. Все більше з'являється таких аномалій. Хоча іноді це вже не аномалії, а – лише звичні пікові (або екстремальні) явища, наприклад, як перенесення пилу Сахари весною у повітрі України [1]. Все це потребує удосконалення методів та засобів системного аналізу стану атмосферного повітря з урахуванням таких аномалій природного та антропогенного походження.

Сучасні інтелектуальні технології аналізу та прогнозування рядів часових даних дають можливість більш гнучко враховувати аномалії, але для цього їх треба спочатку знайти, оцінити характер їх впливу, внести зміни в параметри моделей для опису інших складових цих рядів. Крім того, слід виявити і розкрити відповідні види невизначеності, що мають місце в такій системі.

Загалом, системний аналіз застосовується, як правило, в задачах оптимізації та управління складними системами. В цій сфері такою складною системою є система управління станом атмосферного повітря.

Відповідно до методології системного аналізу треба здійснити декомпозицію складної системи, визначити види невизначеності та запропонувати методи їх розкриття [2-4]. Після цього, варто знайти дані, застосувати до них ці методи і тоді можна ідентифікувати інтелектуальні моделі, здійснити за ними передбачення, провести інтелектуальний аналіз даних, наприклад, за алгоритмом з роботи [5] та розробити рекомендації для управління станом атмосферного повітря, з урахуванням виявлених аномалій.

Багато вчених працюють у цьому напрямку, наприклад, в Європі популярною є декомпозиція цієї системи на основі моделі «DPSIR» [6]. Однак, відсутній дійсно комплексний підхід розв'язання цієї задачі, з розкриттям усіх видів невизначеності на прикладі стану атмосферного повітря в Україні та декомпозиції на основі моделі «DPSIR» з використанням методів системного аналізу та інтелектуальних моделей прогнозування і методів інтелектуального аналізу даних.

Метою роботи є розроблення технології системного аналізу стану атмосферного повітря, з урахуванням аномалій, яка дозволить підвищити точності прогнозування цього стану та, як наслідок, підвищити ефективність функціонування системи управління станом довкілля в заданому регіоні.

Запропонована та описана у роботі технологія була успішно випробувана під час розроблення Програми державного моніторингу в галузі охорони атмосферного повітря на 2023–2027 роки для зони «Вінницька», яка затверджена Міністерство захисту довкілля та природних ресурсів України та Обласною радою та наказом голови Вінницької обласної військової адміністрації (Наказ №722 від 22.05.2023 р.) [7]. Крім того, розроблено декілька програм-ноутбуків на Python на базі платформи Kaggle і викладено у відкритий доступ, що дозволяє усім бажаючим легко використати результати цієї роботи для розв’язання подібних задач для інших регіонів і країн.

Книга містить чотири розділи.

У першому розділі проведено опис стану атмосферного повітря в Україні та системи його моніторингу, охарактеризовано об’єкт дослідження з позиції системного аналізу, наведено аналіз основних підходів та інтелектуальних моделей методів прогнозування часових рядів та пошуку аномалій у цих рядах.

У другому розділі здійснено системний аналіз стану атмосферного повітря з урахуванням впливу аномалій. Описано метод пошуку можливих джерел забруднення та метод оптимізації системи моніторингу повітря.

У третьому розділі наведені методи ідентифікації параметрів сезонності та аномалій інтелектуальних моделей прогнозування стану атмосферного повітря з урахуванням пікових значень.

В четвертому розділі охарактеризовано апробацію результатів дослідження шляхом оптимізації системи моніторингу зони «Вінницька».

Текст книги написаний, переважно, Шмундяком Д. О. за наукового керівництва професора Мокіна В. Б., який здійснював й постановку задач дослідження. Доцент Крижановський Є. М. здійснював прикладне випробування окремих етапів розробленої технології для проєктування системи державного моніторингу в галузі охорони атмосферного повітря на 2023–2027 роки для зони «Вінницька» спільно з ТОВ «ІННОВІННПРОМ» на замовлення Вінницької облдержадміністрації. Це проєктування здійснювалось під час виконання госпдоговірної науково-дослідної роботи на тему «Розробка розділів програми державного моніторингу у галузі охорони атмосферного повітря Вінницької області» (№ держреєстрації 0121U110737, номер у ВНТУ 2859), де Мокін В. Б. був науковим керівником, Крижановський Є. М.

– відповідальним виконавцем, а Шмундяк Д. О. – виконавцем, який відповідав за застосування системного підходу до проведення досліджень. Також, у монографії використано декілька програм-ноутбуків на Python, розроблених Шмундяком Д. О. як індивідуально, так і спільно з Мокіним В. Б. У роботі враховано досвід Мокіна В. Б. та Шмундяка Д. О. як адміністраторів від ВНТУ у сервісі «Кабінет дослідника» від мережі громадського моніторингу EcoCity, де є авторизований доступ до даних більше 30 станцій моніторингу стану атмосферного повітря в межах Вінницької області та міста Вінниця. Мокін В. Б. є одноосібним автором підрозділу 1.2 та 2.1, Крижановський Є. М. є основним автором пунктів 2.3.1, 2.3.2, підрозділів 4.1 і 4.2 та додатків Е, Є, Ж, З, І, які він писав за участі та під науковим керівництвом Мокіна В. Б. Решту матеріалу написано Шмундяком Д. О. за участі та під науковим керівництвом Мокіна В. Б.

Загалом, усі розділи написані з урахуванням порад, доповнень і матеріалів Мокіна В. Б. (1,5 авторських аркушів), а практична апробація описана з урахуванням порад, доповнень і матеріалів Крижановського Є. М. (1,5 авторських аркушів), решта матеріалу написана Шмундяком Д. О.

Відзиви, зауваження і побажання просимо надсилати за адресою: 21021, м. Вінниця, Хмельницьке шосе, 95, кафедра системного аналізу та інформаційних технологій Вінницького національного технічного університету або на пошту sait@vntu.edu.ua.

Автори висловлюють подяку:

– к.т.н., доценту Ящолту Андрію Руслановичу за допомогу у збиранні інформації під час проєктування Програми державного моніторингу в галузі охорони атмосферного повітря на 2023–2027 роки для зони «Вінницька» під час виконання госпдоговірної НДР 2859;

– Копняку Володимиру Євгеновичу, який є третім адміністратором від ВНТУ у сервісі «Кабінет дослідника» мережі громадського моніторингу EcoCity за допомогу з наповнення датасету Kaggle «Air Quality Monitoring from EcoCity» реальними даними весною 2024 року, які використовувались для випробування деяких методів.

РОЗДІЛ 1

ОГЛЯД ТА АНАЛІЗ МЕТОДІВ І ЗАСОБІВ ПРОГНОЗУВАННЯ СТАНУ АТМОСФЕРНОГО ПОВІТРЯ

1.1. Стан атмосферного повітря в Україні та система моніторингу для його аналізу

Атмосферне повітря є одним з компонентів, який, напевно, має найбільший вплив на життєдіяльність людей та усіх інших живих організмів, що проживають на Землі. Відповідно, забруднення повітря провокує досить серйозні наслідки для усієї екосистеми, в тому числі прямо впливає на якість життя людей, збільшення захворюваності, провокує алергічні реакції, тощо. Особливо критично це може впливати на вразливі верстви населення: літні люди, діти, люди з хронічними хворобами. Усе це робить контроль та боротьбу з забрудненням однією з найважливіших проблем сучасності.

Джерела забруднення загалом можна поділити на дві категорії: природні та антропогенні. До переліку природних джерел входять пожежі, пилові бурі, виверження вулканів, тощо. Усе це, безумовно, має суттєвий вплив на якість повітря, проте все ж головним чинником негативної тенденції щодо якості атмосферного повітря є саме людство та його діяльність. Протягом поточного та минулого століття відбувся значний промисловий та технічний розвиток, що створив величезну кількість різного роду джерел забруднення повітря. До цієї категорії можна віднести промислові підприємства, транспорт, використання хімічних засобів в сільському господарстві та побуті тощо. Без належного контролю та політики щодо зменшення викидів забруднювальних речовин, екосистема може досягти критичного стану і перестати бути придатною для життя.

Україна не є винятком – так само мають місце усі ці проблеми і ставиться за мету покращення методів та підходів контролю та зменшення рівня забруднення повітря.

Розглянемо детальніше що саме забруднює атмосферне повітря, як це контролюється та які законодавчі механізми існують для боротьби з цими викликами.

1.1.1. Показники забруднення атмосферного повітря

Забруднення атмосферного повітря, зазвичай, відбувається через викид різних хімічних та біологічних сполук, які спричиняють зміни в атмосфері та чинять цим негативний вплив на живі організми. Основ-

ними забруднювальними речовинами є гази та тверді частинки, які можуть мати як природне, так і антропогенне походження. Деякі забруднюючі речовини можуть мати прямий та швидкий ефект на екосистему, проте часто ефект від забруднення є непрямим і може проявитися та почати впливати на людей та інші організми лише через досить тривалий проміжок часу. Розглянемо основні показники забруднення атмосферного повітря за якими переважно відбувається моніторинг:

1. Тверді частинки (англійською «Particulate Matter», або просто «PM») — дрібний пил, який може проникати в легеневу тканину та кров, що, в свою чергу, може викликати серйозні наслідки для функціонування організму [8]. Його поділяють на декілька груп, в залежності від розміру частинок. Частинки, розміром менше 1 мкм, називаються PM1 (або PM1.0), частинки, розміром менше 2,5 мкм – PM2.5, тверді частинки розміром до 10 мкм, відповідно, – PM10;

2. Озон (хімічна формула «O₃») – газ, який утворюється в результаті реакції сонячного світла та оксиди азоту, що, в свою чергу утворюються в результаті роботи транспортних засобів, об'єктів промисловості, хімічних заводів та електростанцій [9];

3. Діоксид азоту (хімічна формула «NO₂») і оксид азоту (хімічна формула «NO»). Оксиди азоту утворюються в промисловості, шляхом спалювання палива при високій температурі. Також вони можуть проникати в побут із зовнішнього повітря при палінні або через користування газовими плитами. Також, потрапляють у повітря з автомобільних вихлопних газів;

4. Діоксид сірки (хімічна формула «SO₂») – безбарвний газ з різким запахом, який утворюється при спалюванні вугілля або нафти. Джерелами значних обсягів даного газу є побутове опалення, виробництва електроенергії та автотранспорт;

5. Оксид вуглецю (хімічна формула «CO») — безбарвний отруйний газ, який утворюється при спалюванні будь-якого виду палива, що робить його дуже поширеним: ТЕС, ТЕЦ, металургія, нафтопереробна промисловість, транспортні засоби.

1.1.2. Сучасний рівень забруднення атмосферного повітря в Україні

Згідно законодавства та відповідних постанов [10-13], Міністерство захисту довкілля та природних ресурсів щороку готує аналітичний звіт щодо стану навколишнього середовища в Україні. Серед багатьох напрямків, даний звіт в тому числі включає інформації про атмосферне повітря. В умовах воєнного стану, ці звіти публічно не оприлюдню-

ються. Останній публічний звіт характеризує стан природного середовища з 2021 рік. Згідно цього звіту, головними джерелами забруднення повітря є транспорт, підприємства теплоенергетики та переробна промисловість [14].

Річні об'єми кількості забруднювальних речовин, згідно цього звіту, наведено в таблиці 1.1.

Таблиця 1.1 – Обсяги викидів забруднювальних речовин по Україні за 2019-2021 роки [14]

	Викиди забруднювальних речовин, т		
	2019	2020	2021
SO ₂	19803	18169	19643
CO	1255225	1084752	1164873
NO ₂	178778	161781	174420
NO	798	799	896
NH ₃	6	7	8

Як видно з таблиці 1.1, рівень забруднення за основними показниками або залишився на попередньому рівні, або зростає.

Подібно до загально-національного звіту стану навколишнього середовища, щорічно публікуються і регіональні звіти для кожної адміністративної одиниці України. Аналізуючи останній актуальний звіт щодо екологічного стану Вінницької області [15], можна сказати, що рівень забруднення по багатьох показниках загалом має певні позитивні зміни. Основними забруднювальними речовинами повітря є сполуки вуглецю та сірки. Основними ж джерелами, подібно до загальноукраїнського стану, є об'єкти промисловості, сільського господарства та транспорт.

Абсолютні значення кількості забруднювальних речовин дає базове розуміння стану атмосферного повітря, проте, більш показовим є перевищення граничних показників забруднення. В таблиці 1.2 наведено річне середнє значення та максимальне зафіксоване значення перевищення граничних рівнів забруднення для деяких основних показників забруднення атмосферного повітря.

Таблиця 1.2 – Річне середнє значення викидів забруднювальних речовин у Вінницькій області [15]

	Середнє значення забруднювальної речовини за рік	Зафіксоване перевищення граничного рівня забруднення
SO ₂	0,215 мг/м ³	0,54 мг/м ³
CO	2,9 мг/м ³	6,5 мг/м ³
PM	0,36 мг/м ³	4,15 мг/м ³
NH ₃	0,13 мг/м ³	0,36 мг/м ³
CO ₂	0,7571 мг/м ³	-

З таблиці 1.2 можна зробити висновок, що за основними показниками існує значне перевищення граничних показників. Це має негативний вплив на населення та стан регіону та вимагає посилення заходів контролю та зменшення викидів.

1.1.3. Державні механізми контролю та покращення стану атмосферного повітря

Державна політика та системи моніторингу якості атмосферного повітря з історичних причин базується на радянських нормах, проте, протягом останніх років, Україна стрімко рухається в напрямку всебічної інтеграції з Європейським союзом. Серед багатьох норм та моделей, які діють на заході, важливим місцем є контроль за якістю атмосферного повітря і Україна, як офіційний кандидат на вступ до Європейського союзу, зобов'язана адаптуватися до них. Так, наприклад, згідно угоди про асоціацію з ЄС [16] та її додатками, пов'язаними з контролем за навколишнім середовищем, Україна зобов'язується створювати та розвивати системи моніторингу за якістю повітря, запроваджувати плани дій на випадок перевищення різних показників забруднення та зменшувати викиди небезпечних речовин у повітря.

Україна вже має великий прогрес в питанні адаптації до усіх необхідних норм та директив. Європейські підходи щодо моніторингу якості атмосферного повітря вже затверджені на законодавчому рівні переліком постанов та наказів різних державних органів та міністерств [10-12]. Тобто, на сьогодні вже є певна законодавча основа для розвитку мережі державного моніторингу якості атмосферного повітря та певні законодавчі механізми для контролю та боротьбою з джерелами антропогенного забруднення. Суб'єкти господарювання ж зобов'язані слідувати відповідним нормам та скорочувати об'єми шкідливих речовин, які вони викидають в атмосферу.

Згідно закону України [13] передбачається наявність державної системи моніторингу довкілля, в обов'язки якої входить спостережень

за станом і рівнем забруднення навколишнього середовища: якість повітря, стан вод, стан ґрунтів, тощо. Відповідальним за впровадження і розвиток системи моніторингу є Міністерство захисту довкілля та природних ресурсів України. За інформацією міністерства, Державна гідрометеорологічна служба проводить моніторинг забруднення атмосферного повітря у 53 містах України на 162 стаціонарних станціях моніторингу. Крім того, мережа також налічує 2 маршрутних поста спостереження та 2 станції транскордонного переносу. Обов'язкове спостереження проводиться за сімома речовинами: пил, двоокис азоту (NO₂), двоокис сірки (SO₂), оксид вуглецю, формальдегід (H₂CO), свинець та бенз(а)пірен. Деякі з станцій мережі додатково проводять спостереження за декількома додатковими показниками.

Державна екологічна інспекція (Міндовкілля) здійснює відбір проб на джерелах викидів та перевіряє їх на понад 65 параметрів забруднення. Міністерство охорони здоров'я України здійснює спостереження за якістю атмосферного повітря у житловій та рекреаційній зонах, зокрема поблизу основних доріг, санітарно-захисних зон та житлових будинків, на території шкіл, дошкільних установ та медичних закладів в містах та в робочій зоні. Також проводиться аналіз якості повітря у житловій зоні за скаргами мешканців відповідної зони.

На території Вінницької області, дослідження стану та рівню забруднення атмосферного повітря проводиться державною лабораторією Вінницького обласного лабораторного центру Міністерства охорони здоров'я України. Також, на 4 пунктах спостереження Вінницьким центром гідрометеорології ведеться моніторинг атмосферних опадів.

1.1.4. Проєкти громадського моніторингу

Як було згадано в попередніх пунктах, сучасний стан моніторингу якості атмосферного повітря в Україні суттєво покращився протягом останніх років і загалом має тенденцію до подальшого розвитку та модернізації. Проте, це не завжди було так, що давало поштовх до створення та розвитку громадських проєктів, які певним чином намагалися шукати певні рішення щодо контролю та покращення екологічної ситуації в Україні.

Одним з таких проєктів є мережа громадського моніторингу EcoCity. Даний проєкт був створений і продовжує функціонувати з 2019 року за підтримки міжнародної програми «Чисте повітря для України», яка була ініційована громадською організацією «Арніка» [17, 18]. Дана організація займається розробкою та розповсюдженням стан-

цій моніторингу, які можуть бути встановлені компаніями або звичайними людьми для збору інформації про стан атмосферного повітря. Крім того, проєкт розробив інтерактивну мапу забруднення, доступну як веб-додаток, що дозволяє будь кому переглядати інформацію про забруднення повітря майже в режимі реального часу. Станом на початок 2025 року, мережа нараховує понад 350 станцій, розташованих у понад 100 населених пунктах у межах всієї України [18]. Дві такі станції розташовані на території Вінницького національного технічного університету та, відповідно, доступні через згаданий веб-додаток EcoCity. Детальнішу інформацію про можливості станцій моніторингу та про мережу загалом наведено в додатку А.

Окрему увагу варто приділити досить важливому та корисному елементу мережі моніторингу EcoCity – веб-сервіс «Кабінет дослідника» [19]. Як було зазначено раніше, інтерактивна мапа EcoCity дозволяє переглядати актуальну інформацію щодо якості атмосферного повітря, проте не надає можливості перегляду архівів даних, які є корисними для проведення різного роду досліджень. «Кабінет дослідника» вирішує цю проблему, адже сервіс дає можливість користувачам робити запити на отримання архівних даних за усіма наявними показниками за заданий проміжок часу. Доступ до даного сервісу відкривається за умови укладання договору з проєктом EcoCity. ВНТУ уклав цю угоду (договір про приєднання до Програми підтримки дослідників якості повітря України та реєстрацію Кабінету дослідника якості повітря України від 24.03.2023 р), що відкрило доступ до 31-ї станції моніторингу, які розташовані у м. Вінниці та Вінницькій області, з даними за період від 2019 року.

Приклад форми створення запиту на завантаження даних та налаштування часового проміжку за який необхідно завантажити дані наведено на рисунку 1.1 та 1.2 відповідно.

Запит /Request

Архівні дані доступні від 2019-02-20 до 2024-11-11

Оберіть адміністративну одиницю /Select an administrative unit

Вінницька область / Vinnytska oblast

Оберіть параметри моніторингу якості повітря /Select air quality monitoring parameters

[Усі параметри](#) /All parameters

☒	PM1.0 ug/m3	Пил 1.0 мкм мкг/м³
☒	PM2.5 ug/m3	Пил 2.5 мкм мкг/м³
☒	PM10 ug/m3	Пил 10 мкм мкг/м³
☐	Temperature °C	Температура °C
☐	Humidity %	Вологість %
☐	Pressure hPa, Pa	Атмосферний тиск мм.рт.ст.
☐	Noise db	Шум дБ.
☐	RAD uR/h	Радіаційний фон мкР/год
☐	Solar Radiation Wat/m2	Сонячна радіація Вт/м2

Рисунок 1.1 – Форма на створення запиту [19]

Масштаб усереднення результатів /The averaging results scale

20 хвилин / 20 minutes

Дата від /Date from

02/20/2019

Дата до /Date to

11/11/2024

Сформувати запит /Create the request

Рисунок 1.2 – Налаштування часового проміжку та масштабу усереднення даних [19]

Як видно на рисунках 1.1 та 1.2, система дозволяє обрати показники забруднення з списку усіх наявних (важливе уточнення, що даний перелік включає абсолютно усі показники, які теоретично можуть збиратися станціями, проте деякі більш простіші по комплектації станцій збирають дані по обмеженому переліку показників), часовий проміжок за який необхідно отримати набір даних та масштаб усереднення цих даних. Система надає можливість отримати усередненні дані за 20 хвилин, за 1 годину або за 1 добу.

Приклад отриманих даних наведено на рисунку 1.3.

# id_station	datetime	indicator_name	indicator_dimens...	# value
650	2020-01-22 2024-03-11	Temperature 16% Humidity 16% Other (5517) 68%	ug/m3 32% °C 16% Other (4201) 52%	-8.68 102k
650	1/22/2020 00:00:00	NH ₃	ppm	0
650	1/22/2020 00:00:00	NO ₂	ppm	5.2887
650	1/22/2020 00:00:00	CO	ppm	0.002
650	1/23/2020 00:00:00	NH ₃	ppm	0
650	1/23/2020 00:00:00	NO ₂	ppm	7.0416
650	1/23/2020 00:00:00	CO	ppm	0.002
650	1/24/2020 00:00:00	NH ₃	ppm	0
650	1/24/2020 00:00:00	NO ₂	ppm	9.4682

Рисунок 1.3 – Приклад даних EcoCity [19]

Іншим проектом є система моніторингу якості атмосферного повітря YourAirTest. Проект займається збиранням та обробленням інформації про якість атмосферного повітря. Користувачі системи можуть переглядати інформацію в режимі, близькому до реального часу, через відповідний веб-портал (рис. 1.4).

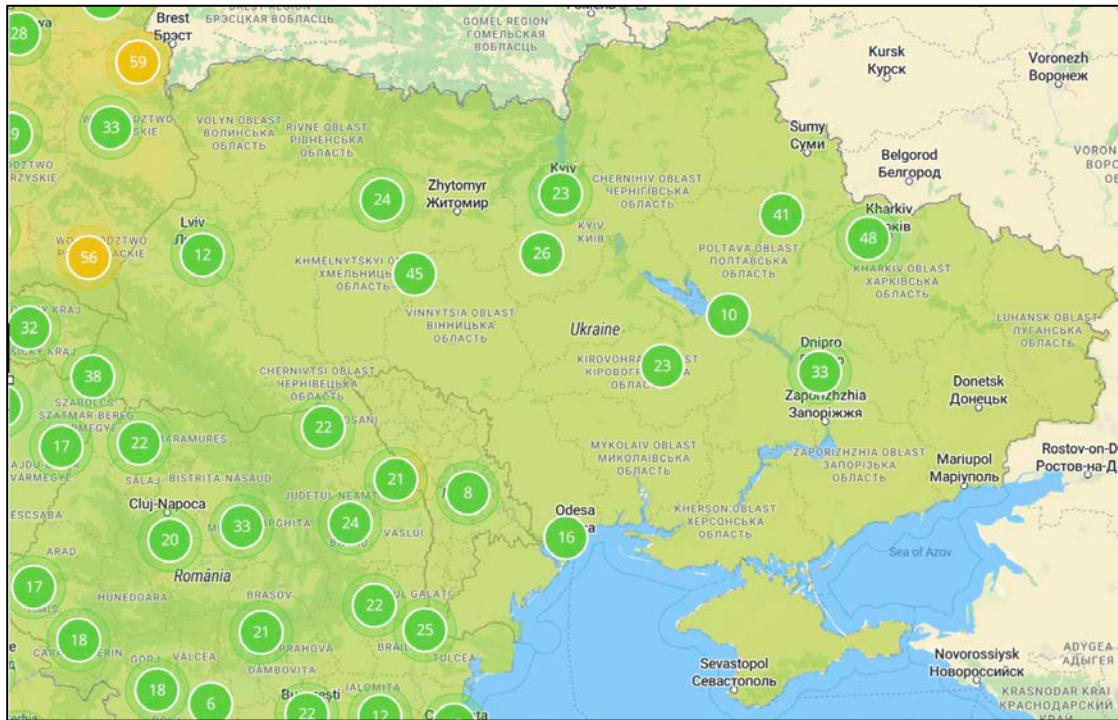


Рисунок 1.4 – Веб-додаток системи «YourAirTest»

Крім того, серед напрямків діяльності системи є аналіз та моделювання пожеж, дослідження та прогнозування погоди, тощо.

Іншим громадським проєктом, який займається екологічними питаннями, в тому числі якістю атмосферного повітря, є SaveEcoBot [20]. Даний проєкт є розробкою громадської організації «SaveDnipro», яка займається аналізом екологічних проблем та пошуком їх рішень. SaveEcoBot – це екологічний чат-бот та інтерактивна веб-мапа, яка дозволяє переглядати дані про забруднення та забруднювачів, а також отримувати сповіщення в разі зміни рівнів забруднення. Інтерактивна мапа містить дані моніторингу якості повітря, надані самою ГО «SaveDnipro», так і містить дані від партнерів даного проєкту. Раніше серед таких партнерів був і проєкт EcoCity, проте, на жаль, в останні роки EcoCity не передає інформацію в SaveEcoBot, відповідно є можливість перегляду тільки застарілих даних і станцій. Детальнішу інформацію про цей проєкт наведено в додатку А.

Варто зазначити, що дані з різних громадських мереж було агреговано Міндовкілля – вони доступні через веб-сервіс «ЕкоЗагроза», приклад вікна якого наведено на рисунку 1.5.

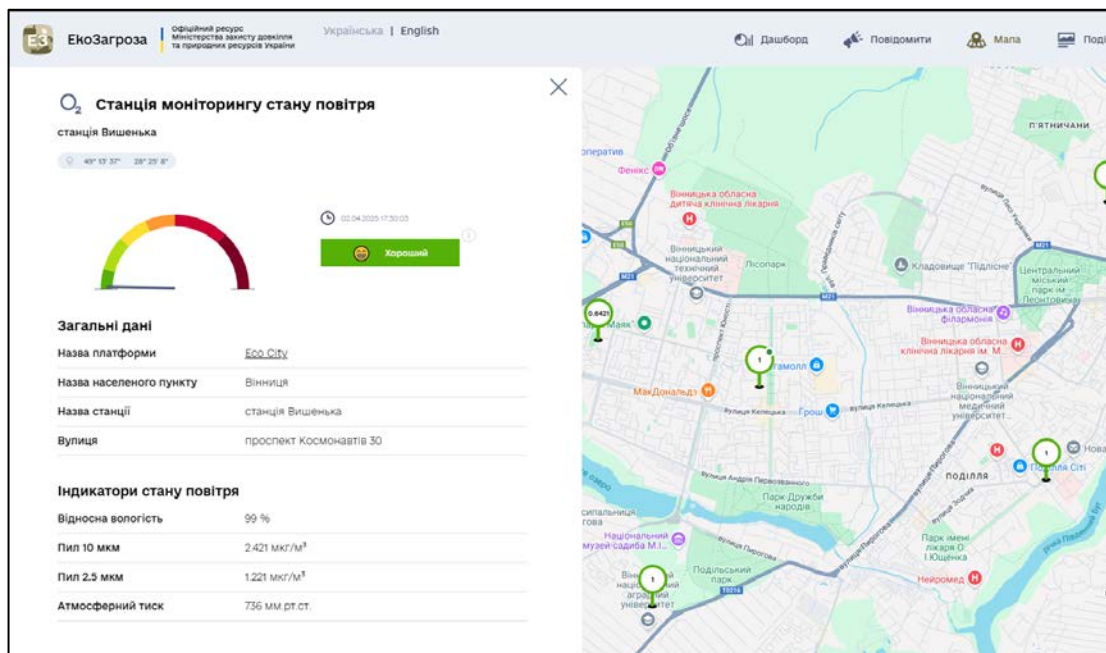


Рисунок 1.5 – Веб-додаток сервісу «ЕкоЗагроза»

1.1.5. Індекс якості повітря

Загальноприйнятим методом інформування населення щодо якості атмосферного повітря є індекс якості повітря (англійською «Air quality index» або скорочено «AQI»). На рівні різних країн визначаються різні категорії та підходи щодо обрахунку та категоризації обрахунків до відповідного рівня забруднення, але, зазвичай, індекс якості повітря це певне значення (деякі країни використовують числові значення, а деякі – відповідними словами), яке відповідає рівню ризику для населення та вразливих його верств (люди з хронічними хворобами, діти, літні люди). До кожного рівня індексу якості повітря також додаються рекомендації для населення, які можуть включати як обмежену, так і повну заборону перебування на відкритому повітрі, оскільки відповідний рівень забруднення може призвести до важких наслідків для здоров'я людини.

Як було зазначено, різні країни мають дещо різні підходи до обрахунку цих значень, але в загальному методологія вимагає збору інформації по певних показниках забруднення та їх усереднення за певний часовий проміжок. В переважній більшості, країни беруть до уваги показники діоксиду сірки (SO₂), великий та дрібних часток пилу (PM₁₀ та PM_{2.5} відповідно), діоксиду азоту (NO₂) та оксиду вуглецю (CO). Потім обраховані значення порівнюються з граничними, що дозволяє віднести поточний стан забруднення до відповідного діапазону.

Серед прикладів існуючих індексів якості повітря можна навести Європейський індекс якості повітря (англійською «European Air Quality

Index»), розроблений Європейським агентством з навколишнього середовища (англійською «European Environment Agency» або скорочено «ЕЕА»), який використовується усіма країнами-членами Європейського союзу [21]. Більш детально про цей індекс наведено в додатку Б.

Також варто згадати і про Український індекс якості повітря, розроблений в рамках міжнародної програми «Чисте повітря для України» і який реалізовано та застосовується мережею громадського моніторингу «EcoCity» [18]. Детальний опис цього методу та критеріїв визначення рівнів забруднення наведено в додатку В.

1.2. Методологія системного аналізу процесів у складних системах

Відомо, що системний аналіз, як наукова методологія, проводиться у 5 основних етапів, які дуже стисло можна подати так (більш розлого див. у [2]):

- 1) постановка задачі;
- 2) виділення об'єкта з навколишнього середовища;
- 3) побудова моделі;
- 4) моделювання, прогнозування та аналіз;
- 5) оптимізація об'єкта або використання результатів аналізу для підтримки прийняття управлінських рішень.

З іншого боку, системний аналіз застосовується, передусім, до складних систем, а складна система – це, як відомо, впорядкована множина структурно взаємопов'язаних і функціонально взаємозалежних систем, які відображають різноманітні й непорівнянні між собою характеристики об'єкта і побудовані для розв'язання багатоцільових задач [4, 22]. Складна система має такі властивості або принципи:

- ієрархічність, тому важливо проводити її декомпозицію на окремі елементи за різними критеріями;
- невизначеність, тому важливо класифікувати усі види невизначеності, які мають місце, і запропонувати підходи чи методи їх розкриття (усунення або мінімізації);
- відкритість до обміну енергією, інформацією чи ресурсами із зовнішнім середовищем, тому варто визначити усі ознаки, які враховують такі взаємовпливи;
- складність зв'язків, яку добре враховують сучасні інтелектуальні моделі, що ідентифікуються з використанням технологій машинного навчання;
- адаптивність, тому є важливим інтелектуальний аналіз системи з передбаченням її поведінки за різними сценаріями з використанням сучасних технологій штучного інтелекту.

Отже, проведення системного аналізу тісно пов'язано із проведенням досліджень з використанням різних методів науки про дані (англ. «Data Science») та технологій штучного інтелекту, а такі дослідження, у свою чергу, проводяться за таким алгоритмом [5, 23]:

1. Поставити задачу.
2. Сформувати набір даних.
3. Здійснити попереднє оброблення даних, провести розвідувальний аналіз даних та застосувати оброблення ознак на вимогу цього аналізу, ще раз провести розвідувальний аналіз даних, на основі якого висунути гіпотези про інтелектуальні моделі, які варто побудувати для розв'язання поставленої задачі.
4. Побудувати інтелектуальні моделі та вибрати серед них оптимальну.
5. Провести на основі оптимальної моделі інтелектуальний аналіз даних, зробити висновки та рекомендації щодо об'єкту дослідження, можливо, повторити якісь етапи.

Щоб поєднати усі ці методології в єдине ціле, пропонується інтелектуальний системний аналіз довільної складної системи з використанням методів науки про дані (англійською «Data Science») та технологій штучного інтелекту:

1) *постановка задачі*: сформулювати поняття складної системи та охарактеризувати її в цілому, здійснити декомпозицію за різними критеріями на підсистеми та елементи, стисло їх проаналізувати (їх стан, особливості виділення із навколишнього середовища та взаємодію з ним і між собою), на основі чого – класифікувати види невизначеності, які мають місце і потребують розкриття, визначити критерії оптимізації та види метрики для вибору оптимальної моделі для аналізу даних, сформулювати обмеження, які мають місце, та здійснити класичну математичну постановку задачі системного аналізу;

2) *виділення об'єкта з навколишнього середовища та формування набору даних*: запропонувати і застосувати методи та підходи розкриття невизначеності усіх підсистем та елементів і процесів у них, визначених на попередньому етапі, зокрема й для виділення об'єкта з навколишнього середовища; знайти чи сформулювати набір даних для аналізу; провести його розвідувальний аналіз, можливо, знайти додаткові види невизначеності, які властиві системі в цілому чи її окремим підсистемам, застосувати оброблення ознак на вимогу цього аналізу (пошук аномалій, усунення помилкових даних, балансування елементів набору, інтерполювання даних тощо), потім ще раз провести розвідувальний аналіз даних, результатом якого має бути чіткий алгоритм подальших дій щодо розв'язання задачі з використанням технологій штучного інтелекту;

3) *моделювання процесів у складній системі, пошук закономірностей, передбачення і прогнозування даних*: вибір архітектури інтелектуальних моделей, розроблення методів і підходів ідентифікації їх параметрів, побудова інтелектуальних моделей та вибір оптимальної моделі серед них, можливо, з використанням оригінального критерію оптимальності (метрики); її застосування, аналіз важливості ознак моделі, після чого, можливо, удосконалення моделей і повторення усіх цих операцій;

4) *аналіз даних* на основі застосування побудованої на попередньому етапі оптимальної моделі: аналіз важливості ознак моделі, аналіз так званих «викидів» (чи помилок 1 і 2 роду), аналіз і візуалізація виявлених закономірностей, після цього можливим є повторення попереднього чи усіх попередніх етапів або – формування висновків щодо складної системи та даних про неї;

5) *оптимізація складної системи* як системи управління або обґрунтування множини управлінських рішень щодо неї, або оптимізація інтерактивної людино-машинної системи взаємодії, наприклад чат-боту на основі великих мовних моделей тощо.

Етапи 2-4 або 3 і 4 можуть застосовуватись багато разів для різних аспектів системи, які є важливими на етапі 5.

Пропонуємо проводити розв’язання поставленої задачі, з урахуванням цієї концепції. А тому наступні підрозділи і пункти цього розділу будуть присвячені етапу постановки задачі. Буде сформульовано поняття складної системи та наведена характеристика її в цілому, здійснена декомпозиція за різними критеріями на різні складові та їх аналіз, на основі чого – класифіковані види невизначеності, які мають місце і потребують розкриття, визначені критерії оптимізації та види метрики для вибору оптимальної моделі для аналізу даних, сформульовано обмеження, які мають місце, та здійснена класична математична постановка задачі системного аналізу.

У наступних підрозділах першого розділу буде проведено аналіз відомих моделей і підходів розв’язання поставленої задачі з використанням технологій штучного інтелекту. Також буде систематизовано стан розкриття виділених видів невизначеності для складових за усіма виділеними видами декомпозиції, що дозволить з’ясувати які методи та підходи їх розкриття ще слід розробити чи удосконалити. А тоді буде поставлено більш конкретні задачі дослідження, які будуть розв’язані у наступних розділах.

Можна виділити такі види декомпозиції за різними критеріями [5]:

– *функціональна* – за функціями чи призначенням елементів (UML діаграма варіантів використання або просто схема з різними блоками тощо);

- *структурна* або *компонентна* – поділ на окремі складові за фізичним чи логічним принципом або поділ на класи (UML діаграма класів) чи компоненти (UML діаграма компонентів) тощо;
- *ієрархічна* – за рівнями організації чи впливу (об’єктовий, місцевий, локальний, регіональний, національний, глобальний або UML діаграма пакетів);
- *геометрична* або *просторова* – за просторовими зв’язками (моделі у 2D/3D тощо) або за географічними координатами з розбиттям на полігони (зони), лінії, з формуванням геометричних мереж тощо;
- *часова* або *хронологічна* – урахування послідовності етапів у часі та динаміки функціонування, UML діаграма часових характеристик;
- *процесна* – за стадіями чи етапами роботи системи (підготовка, запуск, супровід, завершення, візуалізація результатів);
- *факторна* – за ключовими впливовими чинниками (екологічні, економічні; зовнішні, внутрішні та інші підходи щодо класифікації);
- *галузева* або *предметна* – за галузями або предметними областями (інформаційні технології, транспорт, медицина, екологія, зв’язок, енергетика, промисловість);
- *поведінкова* – за процесами, потоками дій або змінами станів системи (UML діаграма діяльності, станів);
- *взаємодійна* або *сценарна* – за сценаріями взаємодії об’єктів чи підсистем (UML діаграма взаємодії, моделі сценаріїв);
- *інформаційна* – за інформаційними потоками (інформаційні системи і моделі, моделі комунікацій);
- *кібернетична* – на основі кібернетичної моделі технічної системи з виділенням складових, які відповідають за вхід, стан системи, збурення, завади, вихід у різних режимах.

Більшість задач системного аналізу містять такі основні види невизначеності [4]:

- *ситуаційна* – відсутність точних знань про ймовірні ситуації і стан складної системи, обумовлена невизначеністю впливу неконтрольованих факторів на неї у процесі практичної діяльності або неповне чи неправильне розуміння поточного стану системи та її оточення; причини: недостатність даних, складність або швидка зміна ситуації, проблеми моніторингу;
- *цільова* – невизначеність вибору і досягнення цілей у багатокритеріальних задачах прийняття рішень або відсутність чітко визначених, погоджених чи вимірюваних цілей, пріоритетів управління для системи; причини: конфлікти інтересів, багатоваріантність рішень або неоднозначність пріоритетів;

- *інформаційна* – неповнота, неточність, суперечливість або ненадійність доступної інформації; причини: обмеженість даних, проблеми моніторингу або інформаційні бар'єри;
- *інформаційна невизначеність конфліктів* – невизначеність вибору задумів і планів у процесі взаємодії партнерів або протидії конкурентів чи супротивників; причини: обмеженість інформації про варіанти та ймовірнісні характеристики щодо вибору задумів, планів партнерів, конкурентів чи супротивників;
- *структурна* – неповне знання про елементи системи або неточність уявлень про зв'язки між ними; причини: обмеженість інформації про систему та зв'язки у ній, проблеми моніторингу;
- *параметрична* – невизначеність значень параметрів моделі; причини: обмеженість даних або їх варіативність, проблеми моніторингу;
- *динамічна* – непередбачувані зміни у поведінці системи через нелінійність або затримки у зворотних зв'язках;
- *стохастична* – випадкові флуктуації чи ймовірнісна природа процесів у системі;
- *факторна* – неповна інформація про вплив зовнішніх факторів: навколишнього середовища та антропогенних чинників, макро- та мікроекономіки, законодавства тощо;
- *модельна* – спрощення або невизначеність щодо варіантів моделей складної системи;
- *емерджентна* – непередбачувані властивості, які виникають через взаємодію елементів системи.

Другий розділ буде присвячений методам розкриття невизначеності усіх підсистем та елементів складної системи і процесів у них, визначених у першому розділі, зокрема й для виділення об'єкта з навколишнього середовища, побудови, розвідувального аналізу даних (англійською «Exploratory Data Analysis» або скорочено «EDA»), обробленню ознак (англійською «Feature Engineering» або скорочено «FE») та формуванню алгоритму подальшого розв'язання задачі з використанням технологій штучного інтелекту та інтелектуального аналізу даних.

У третьому розділі будуть наведені розроблені чи удосконалені методи ідентифікації параметрів інтелектуальних моделей з використанням методів машинного навчання та — алгоритми і засоби інтелектуального аналізу даних та рекомендації щодо використання результатів його для оптимізації складних систем чи для підтримки прийняття рішень.

У четвертому розділі будуть наведені результати випробувань розроблених моделей, методів, алгоритмів і програм для розв'язання реалі-

льних задач, наведені результати, охарактеризоване можливе чи реальне впровадження, запропоновані рекомендації щодо подальших досліджень.

Така підхід щодо опису методології системного аналізу та результатів її застосування є універсальним – рекомендується його застосовувати для різних задач.

1.3. Характеристика об'єкта досліджень з позицій системного аналізу

Охарактеризуємо об'єкт досліджень з використанням наведеної у підрозділі 1.2 методології системного аналізу. Для цього, спочатку сформулюємо поняття складної системи у цій задачі.

1.3.1. Формалізація системи управління стану атмосферного повітря як складної системи

В задачі управління станом довкілля в цілому, та атмосферного повітря зокрема, в ЄС основою екологічної політики є так звана модель «DPSIR» [6] (рис. 1.6).

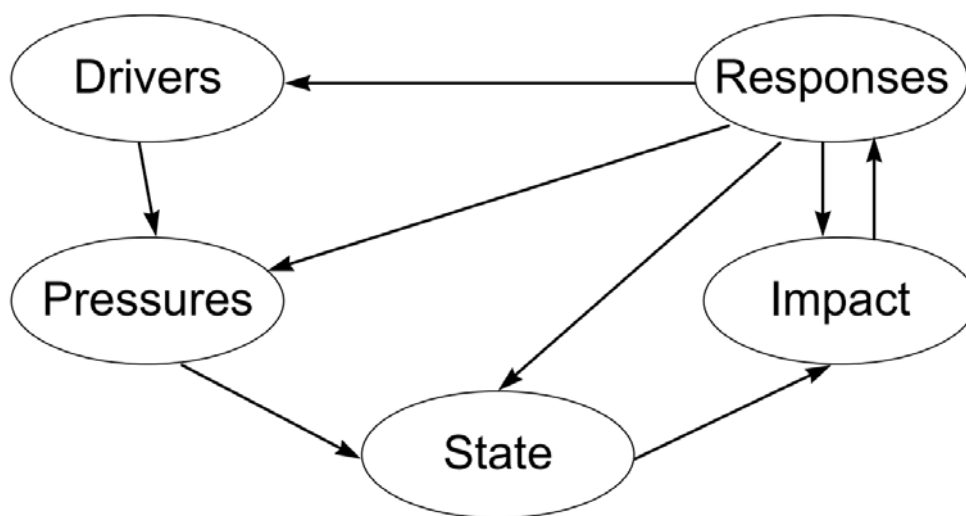


Рисунок 1.6 – Структурна схема моделі «DPSIR»

Як видно, з рис. 1.6, прийняття рішень (англ. «Responses» — «R») стосується [6]:

- вибору місць розташування та вибору складу продуктивних сил (заводів, електростанцій, сільськогосподарських полів тощо) (англ. «Drivers» — «D»);
- вибору місць розташування, очисних систем та способів викидів газів та обґрунтування гранично допустимих дозволів щодо їх обсягу

та допустимих забруднювальних речовин тощо, які характеризують безпосередній тиск продуктивних сил на довкілля (англ. «Pressures» — «P»);

- вибору місць розташування та параметрів систем моніторингу стану довкілля (англ. «State» — «S»);

- вибору рекомендацій для органів влади, екоінспекцій та інших контрольних установ для підвищення ефективності та оперативності контролю, рекомендацій для населення щодо зменшення негативного впливу на них поганого стану довкілля (щільно закривати вікна, коли палять листя або прилітає пил із Сахари тощо) (англ. «Impact» — «I»).

Як було зазначено вище, на відміну від класичного підходу, ця робота присвячена системному аналізу системи, яка зосереджена, в першу чергу, на врахуванні аномалій або пікових значень показників у системі.

Здійснимо декомпозицію цієї системи за різними критеріями.

1.3.2. Декомпозиція складної системи за різними критеріями

Пропонується застосувати такі види декомпозиції системи управління станом атмосферного повітря регіону з урахуванням, передусім, впливу аномалій чи пікових значень показників або параметрів:

- *функціональна* (виділення підсистем за функціями чи їх призначенням): D_{DPSIR} (D_D – джерела-забруднювачі, D_P – види навантаження, D_S – стан системи, D_I – наслідки, D_R – реакція);

- *просторову* (за просторовими зв'язками та географічними координатами): D_G – зон (Вінницька область), агломерацій (м. Вінниця або агломерація «Вінницька»), локальну (мікрорайон Вишенька), організацій (ВНТУ) – у цій роботі основна увага буде приділятися рівню зон (областей) D_{G-O} ;

- *часову*, оскільки різні аномалії можуть бути різними у різні інтервали часу: $D_T(\Delta T)$ – характеристика різних підсистем та вибірки спостережень різних показників з узагальненням за інтервал часу ΔT – 1 хв., 20 хв., годину, добу, тиждень, місяць, рік тощо;

- *сезонну*, оскільки повторюваність інтервалів часу дозволяє зрозуміти це – аномалія чи лише сезонне пікове значення: $D_{SEC}(\Theta)$ – з сезоном, тривалість Θ – добова, тижнева, місячна, квартальна, річна тощо.

Інші види декомпозиції мають менше значення для даної задачі з акцентуванням на вплив аномалій.

Виділимо та охарактеризуємо види невизначеності, які мають місце для цих складових.

1.3.3. Класифікація видів невизначеності

Виділимо види невизначеності, які мають місце у складній системі управління станом атмосферного повітря регіону з урахуванням, передусім, впливу аномалій чи пікових значень показників або параметрів.

Проаналізуємо види невизначеності для складових за декомпозицією D_{DPSIR} :

- D_D – джерела-забруднювачі:

- *інформаційна* (неповна чи неточна інформація про місця найбільшої щільності розташування підприємств-забруднювачів – U_{I1} або про найбільші обсяги гранично допустимих значень забруднювальних речовин у викидах – U_{I2});

- D_P – види навантаження:

- *інформаційна* (неповна чи неточна інформація про місця розташування джерел викидів та про фактичні викиди – U_{I3}),

- *ситуаційна* (відсутність достовірної інформації про перелік усіх забруднювачів, через те, що якесь підприємство здійснює викид, але не подає про це звітність, або фальсифікує інформацію про реальні викиди) – U_{S1} ;

- D_S – стан системи:

- *інформаційна* (неточні дані про значення показників якості атмосферного повітря – U_{I4}),

- *цільова* для системи збирання даних моніторингу (невизначеність вибору місць розташування існуючих чи нових постів моніторингу) – U_{T1} ;

- *цільова* для системи комплексного оцінювання якості атмосферного повітря (невизначеність щодо критеріїв якості атмосферного повітря в цілому) – U_{T2} ;

- *цільова* для аналізу аномальних та пікових значень показників *якості* атмосферного повітря (як правило, аномалії фільтрують за доданням порогу – має місце невизначеність щодо вибору варіантів цього порогу) – U_{T3} ;

- *модельна* (невизначеність щодо варіантів моделей складної системи) – U_{M1} ,

- *параметрична* (невизначеність щодо значень параметрів моделей) – U_{P1} ;

- D_I – наслідки або вплив:

- *структурна* (неповне знання про елементи системи або неточність уявлень про зв'язки між ними) – U_{ST1} ;
- *сценарна* (відсутність чи неточність прогнозів динаміки системи для різних сценаріїв розвитку) – U_{SC1} ;
- D_R – реакція або прийняття управлінських рішень:
- *інформаційна* (неповна чи неточна інформація про вже прийняті рішення – U_{I5});
- *цільова* (невизначеність вибору рішень) – U_{T4} .

Види невизначеності для складової D_{G-O} , тобто для системи управління станом атмосферного повітря зони (наприклад, адміністративної області України) на додаток до наведених вище:

- *структурна* (неповне знання про взаємозв'язок між різними показниками стану атмосферного повітря на різних постах спостережень – наявність цих знань дозволила б за даними сусідніх постів здійснювати прогнозування показників одних по інших, що компенсувало б інші види невизначеності, передусім інформаційну) – U_{ST2} .

Види невизначеності для складових $D_T(\Delta T)$ на додаток до наведених вище:

- *ситуаційна* (у різні інтервали часу можуть бути різні аномалії чи пікові значення, які важко виявляти) – U_{S2} ,
- *цільова* (як правило, аномалії фільтрують за доланням порогу – має місце невизначеність щодо вибору варіантів цього порогу) – U_{T5} .

Види невизначеності для складових $D_{SEC}(\Theta)$ на додаток до наведених вище:

- *цільова* (вибір яка саме сезонна складова або їх сукупність мають місце) – U_{T6} ;
- *параметрична* (невизначеність щодо параметрів моделей, які описують вплив сезонної складової: період, вікно впливу, коефіцієнт регуляризації або сила впливу) – U_{P2} .

1.3.4. Підходи та методи щодо розкриття видів невизначеності

Охарактеризуємо яким чином можна розкрити (максимально мінімізувати чи й усунути) види невизначеності, виділені у попередньому пункті, з урахуванням наведеної у підрозділі 1.1 інформації.

Для розкриття інформаційної невизначеності про джерела-забруднювачі D_D та види навантаження D_P , достатньо подати запит до Державної служби статистики України та органів влади щодо інформації із форми 2ТП «Повітря» про джерела забруднення та інформацію про ро-

зташування їх джерел викиду. На жаль, там буде тільки офіційна інформація, яка потребує уточнення щодо незареєстрованих джерел забруднення або уточнення фактичних викидів. Для цього варто розробити метод, який дозволить оцінювати можливий вплив додаткових джерел забруднення та здійснити їх локалізацію у просторі та часі, точніше – визначити ділянку та в які роки (чи місяці) їх додатковий вплив був відчутним. Це дозволить розкрити ситуаційну невизначеність щодо D_p .

Щодо розкриття інформаційної невизначеності для D_s , достатньо отримати інформацію з систем моніторингу. Краще використовувати інформацію з державної системи моніторингу, але, за відсутності такої, у першому наближенні, можна використовувати й дані громадської системи моніторингу, які в Україні доступні на вище згаданих порталах SaveEcoBot та EcoCity. Однак, цієї інформації може бути недостатньо, тоді варто розширяти мережу, передусім державного моніторингу, з виведенням інформації на порталах державних органів влади та місцевого самоврядування. Такі пости (станції) є доволі дорогими, тому потрібна система критеріїв для визначення їх місць розташування, показників спостережень, регулярності вимірювань та пріоритетності їх монтажу. Для цього потрібно розробити підхід формування такої системи критеріїв, орієнтовану на аномальні та максимальні показники складових моделі DPSIR: де вже є чи будується найбільша кількість (щільність у просторі) джерел забруднення, де найбільше проживає населення, зокрема й вразливого, де мають місце найбільші викиди тощо. Це дозволить розкрити й цільову невизначеність для системи збирання даних моніторингу для складової D_s . Застосування такої системи критеріїв тісно пов'язано з розкриттям інших двох підвидів цільової невизначеності для цієї складової:

1) цільова невизначеність для системи комплексного оцінювання якості атмосферного повітря – пропонується як комплексну оцінку використовувати загально прийнятий в ЄС індекс атмосферного повітря (див. 1.1.5);

2) цільова невизначеність для аналізу аномальних та пікових значень показників якості атмосферного повітря – пропонується використовувати директиву 2008/50/ЄС європейського парламенту та ради [24] та постанову Кабінету Міністрів України №827 [10] для аналізу найбільших та середніх значень, але треба систематизувати цей підхід.

Для розкриття модельної невизначеності щодо складової D_s , тобто побудови оптимальної моделі для моделювання і прогнозування даних моніторингу стану атмосферного повітря, потрібно провести

огляд відомих моделей та вибрати множину тих, які доцільно застосовувати в даній задачі. Особливу увагу слід приділити питанням розкриття параметричної невизначеності щодо врахування аномалій, визначення видів та параметрів складових для урахування сезонності цих моделей. Причому, під час розкриття цієї параметричної невизначеності важливо приділити увагу розкриттю ще й таких видів невизначеності:

1) ситуаційної та цільової невизначеності для складових $D_T(\Delta T)$ щодо того як виявляти аномалії та пікові значення на різних інтервалах (і на яких саме) часу;

2) цільової та параметричної невизначеності для складових $D_{SEC}(\Theta)$ щодо того яка саме сезонна складова або їх сукупність мають місце та які їх параметри.

З метою кращого розуміння підходів щодо розкриття усіх цих видів невизначеності, пов'язаних з параметрично невизначеністю, варто вибрати регіон для аналізу, знайти для нього дані та провести їх розвідувальний аналіз, на якому відпрацювати усі методи і підходи.

Для більш повного розкриття сценарної невизначеності для складової D_I (та для складової D_{G-O} в цілому) потрібно більше інформації про усі взаємозв'язки в системі, але, як правило, така інформація є недостатньою. Мережа моніторингу в Україні, на жаль, досі не є достатньо розвиненою. Крім того, метеорологічні показники, як правило, не є відкритими даними і спостерігаються теж не дуже точно, через брак фінансування всієї галузі. А тому недостатність інформації можна спробувати компенсувати використанням інтелектуальних моделей та засобів для аналізу даних складової D_S і для прогнозування окремих показників якості атмосферного повітря, і для оцінювання їх взаємовпливу, наприклад на основі байєсівського моделювання. Наявність моделей для прогнозування стану атмосферного повітря окремих регіонів та моделей для оцінювання їх сукупного впливу на інші регіони дозволить програвати сценарії зміни стану в цілому, що дозволить розкрити сценарну невизначеність для складової D_I .

Запит до органів влади дозволить розкрити інформаційну невизначеність щодо складової D_R про вже прийняті управлінські рішення, відображені у різного роду звітах, оглядах, прес-релізах. Цьому може допомогти й аналіз матеріалу на сайтах та сторінках у соцмережах органів влади і місцевого самоврядування, які опікуються екологічною політикою, зокрема й у сфері екоінспекційного контролю. А технології сучасного генеративного штучного інтелекту дозволять на основі цієї текстової інформації сформулювати множину можливих управлінських рішень, серед яких можна налагодити вибір одним із відомих методів

оптимізації. Це все дозволить розкрити й цільову невизначеність щодо складової D_R .

Отже, для розкриття усіх цих видів невизначеності необхідно здійснити наступне:

1) провести огляд інтелектуальних моделей, методів і засобів прогнозування даних часових рядів та вибрати множину тих, серед яких варто шукати оптимальні;

2) провести огляд методів і засобів автоматизованого пошуку аномалій та пікових значень у часових рядах;

3) запропонувати метод пошуку можливих незареєстрованих джерел викидів за даними точкового моніторингу стану атмосферного повітря;

4) розробити новий підхід до формування системи критеріїв для вибору параметрів місць нових постів спостережень державної системи моніторингу стану атмосферного повітря регіону та визначення пріоритетності їх монтажу, з урахуванням аномалій та максимальних значень різних показників, які будуть братись до уваги;

5) розробити методи визначення параметрів інтелектуальних моделей для прогнозування даних про стан атмосферного повітря регіону, з урахуванням аномалій та сезонних особливостей;

6) удосконалити метод аналізу взаємозв'язків між різними показниками стану атмосферного повітря різних регіонів на основі інтелектуальних моделей;

7) вибрати регіон для аналізу, знайти для нього реальні дані про стан атмосферного повітря (бажано дані і державного, і громадського моніторингу) та провести їх розвідувальний аналіз, а потім на ньому відпрацювати усі методи і підходи.

1.4. Огляд інтелектуальних моделей, методів і засобів прогнозування даних часових рядів

Задачі аналізу та прогнозування часових рядів набули широкого поширення та мають практичне застосування у багатьох сферах. Часто важливим є не просто побудова моделі, здатної точно і надійно прогнозувати ряди даних. Цінним може бути і сама модель, а особливо ідентифікований період і частота сезонності. Ці та інші параметри моделі можуть сказати багато корисної інформації про сам ряд. Іноді важливо припасувати модель не до останніх значень ряду, а до усіх значень за весь період спостережень, щоб краще зрозуміти відхилення та зміну його поведінки у різний час. Всі ці задачі вирішуються з використанням технологій аналізу і прогнозування, оснований на специфічних моделях

часових рядів, які враховують час, коли було зафіксоване кожне значення.

Розвідувальний аналіз часових рядів іноді ще називають діагностуванням часового ряду. Цей процес містить усі ті ж етапи, які і класичні задачі машинного навчання і пов'язаний з пошуком та усуненням пропущених даних, виявленням аномалій, пошук чи синтез нових ознак, аналіз їх важливості, аналіз закономірностей самого ряду. Однак, в силу залежності в часі та статистичної природи самих моделей часових рядів, ці етапи мають чимало відмінностей.

Як правило, аналіз часового ряду містить такі етапи [5]:

1. Аналіз пропущених значень і регулярності кроку. Передбачає пошук пропусків в даних та заповнення цих пропусків певними очікуваними даними, обрахованих за рахунок аналізу наявних. Деякі моделі для роботи з часовими рядами є менш вибагливими до пропусків, тому цей крок не є обов'язковим і може навпаки просто псувати оригінальні дані без суттєвого приросту ефективності роботи відповідних моделей.

2. Аналіз аномалій ряду. Є важливим кроком, оскільки наявність аномалій зазвичай суттєво погіршують точність прогнозів, зроблених моделями машинного навчання. Детальніше про цей етап буде згадано в наступних розділах.

3. Аналіз закону розподілу ряду. Багато методів аналізу даних побудовано, виходячи з гіпотези про нормальність закону розподілу цих даних. Тому важливо перевіряти цю гіпотезу і якщо вона не виконується, то потрібно спробувати нормалізувати дані. Поширеною проблемою для даних спостережень за станом довкілля є ситуація коли більш частими є менші значення. В такому випадку можна спробувати застосувати математичну трансформацію і досягти стану ближчого до нормального закону розподілу.

4. Аналіз стаціонарності ряду. Стаціонарним є ряд, в якому статистичні показники незмінні в часі. А отже, статистичні показники різних вибірок ряду будуть однаковими чи зіставними у межах заданої похибки. Класичним прикладом нестационарного ряду є температура повітря в Україні протягом року. Якщо взяти вибірку взимку чи влітку, звичайно, середнє буде суттєво відрізнятись. За можливості, ряд варто стаціонаризувати, що буде особливо корисно при використанні нейромережових моделей.

5. Аналіз сезонності ряду. Головною відмінністю моделей часових рядів від інших моделей машинного навчання є те, що вони спеціалізуються на врахуванні сезонності рядів. Сезонність, як правило, описується рядами Фур'є [5]:

$$y(t) = \frac{a_0}{2} + \sum_{i=1}^n a_i \cos\left(\frac{2\pi}{T}it\right) + \sum_{i=1}^n b_i \sin\left(\frac{2\pi}{T}it\right). \quad (1.1)$$

Сучасні функції для навчання моделей часових рядів, як правило, самі ідентифікують параметри Фур'є. Для цього їм потрібно лише 2 гіперпараметри: період T і кількість гармонік ряду Фур'є n .

Результати усіх застосованих етапів розвідувального аналізу застосовуються для прийняти рішення про те, які далі будувати моделі та які саме параметри відповідного набору даних використовувати. Процес вибору значущих параметрів називають Інженерія ознак (FE). У разі, якщо далі планується пробувати використовувати класичні багатofакторні моделі машинного навчання (регресійні, дерева рішень, звичайні нейромережі тощо), то потрібно спробувати знайти або синтезувати побільше ознак, навіть, якщо задано тільки один фічер. Існують Python-бібліотеки, які дозволяють синтезувати велику кількість фічерів для заданого часового ряду. Обчислюється велика кількість різних статистичних характеристик заданого часового ряду (не тільки дисперсія чи середньоквадратичне відхилення, а й енергія та ін.) із різним «вікном».

Безпосередньо для прогнозування часових рядів можна використовувати нейромережеві рекурентні моделі, зокрема GRU та LSTM. Але, частіше, використовуються специфічні для часових рядів моделі ARIMA та Prophet.

Математичний апарат ARIMA (англ. «Autoregressive Integrated Moving Average» – модель авторегресії та проінтегрованого ковзного середнього (АРПКС)) детально описаний у підручнику одного зі спів-авторів [5]. Для її автоматизації на Python використовується один із двох варіантів.

Математично модель Prophet для моделювання та прогнозування значень ряду $y(t)$, залежно від часу t , записується таким чином [5]:

- для адитивного випадку

$$y(t) = g(t) + s(t) + h(t) + \varepsilon_t, \quad (1.2)$$

- для мультиплікативного випадку:

$$y(t) = g(t) \cdot s(t) \cdot h(t) \cdot \varepsilon_t, \quad (1.3)$$

де $g(t)$ – тренд ряду (логістична або кусково-лінійна апроксимація даних), $s(t)$ – сезонна складова, апроксимована рядом Фур'є, $h(t)$ – складова, яка враховує вплив свят чи інших аномалій, які відбуваються нерегулярно протягом одного або кількох днів і діють з певним «вікном», тобто, в діапазоні певних дат, ε_t – похибка («шум») з нульовим середнім, розподілена за нормальним законом.

Модель Prophet враховує аномалії даних та ефективно працює з пропущеними даними (для даних моніторингу, за умов відключення світла чи Інтернету, це – актуальна проблема). З власного досвіду авторів відомо, що модель може досить ефективно працювати з параметрами за замовчуванням [25], але все ж кращої ефективності можна досягти за рахунок використання багатьох параметрів, передусім, про аномалії даних та щодо опису сезонних складових.

Задачі прогнозування стану атмосферного повітря з урахуванням аномалій має деякі особливості:

- можуть мати місце локальні у часі аномалії і треба вміти їх ідентифікувати;
- часовий ряд може мати декілька видів сезонності і треба вміти ідентифікувати їх період та параметри для їх опису.

Урахування цих особливостей дозволить суттєво підвищити адекватність моделей часових рядів та точність прогнозування за ними.

1.5. Огляд методів і засобів автоматизованого пошуку аномалій в часових рядах

Як було зазначено в попередньому розділі, одним з етапів аналізу часових даних є пошук аномальних значень – подій або елементів, які не відповідають певним шаблонам поведінки або іншим елементам даного набору. Процес виявлення аномалій є фундаментальною дослідницькою проблемою, якою займаються науковці різних країн, відповідно, не дивним є те, що наукова література надає досить широкий перелік існуючих методів та підходів [26].

З власного багаторічного досвіду роботи з даними стану різних природних середовищ, можна сказати що є такі проблеми, пов'язані з впливом аномальних і пікових значень окремих показників стану атмосферного повітря на значення інших показників:

1. Показники стану довкілля зазвичай мають певну циклічність, обумовлену впливом сонячної активності (фотосинтез), забрудненням (у воді - скидання стічних вод, у повітрі – трафіком автотранспорту вдень, особливо у години пік) тощо. Ця циклічність може бути різною у різний час, але спільною є форма сигналу: мінімум, наростання, максимум, спадання, мінімум, що відповідає півхвилі.

2. Моделі для роботи з часовими рядами сильно залежать від наявності аномальних значень та параметрів різних видів сезонності. Параметри сезонності слід задавати – вони не визначаються автоматично. А для періодичних рядів важливо знати, як мінімум, період ряду і порядок ряду Фур'є. Складність пошуку одночасно параметрів аномалій та сезонності ряду ще полягає і в тому, що це – взаємозалежні задачі, а

отже аномалії у різних періодах ряду теж будуть різні. А для цього треба знати ці періоди. Однак, щоб знайти період ряду, слід спочатку відфільтрувати аномалії. Особливо задача ускладнюється, якщо має місце декілька різних видів сезонності одночасно або сезонності поєднуються мультиплікативно.

3. Даним моніторингу якості повітря властива зміна періодичності спостережень та наявність пропущених даних.

Розглянемо декілька відомих методів пошуку аномалій та як можна автоматизувати процес пошуку за рахунок використання мови програмування Python.

1.5.1. Аналіз відомих методів пошуку аномалій

Популярним методом пошуку аномалій в задачах аналізу часових рядів є Isolation Forest. Isolation Forest – це алгоритм виявлення аномалій, який працює шляхом ізоляції аномалій у даних [27]. Алгоритм роботи даного методу виглядає наступним чином:

1) з вхідного набору даних випадковим чином обирається підвибірка;

2) обирається випадкова ознака та поріг – випадкове значення між мінімальними та максимальними значеннями обраної ознаки;

3) точка набору даних порівнюється з пороговим значенням. Якщо воно менше порогу, то точка додається у ліву гілку дерева, якщо ж більше – у праву;

4) повторюється крок 3 поки усі точки набору даних не будуть розділені в окремих гілках дерева або поки не досягнуто заданої максимальної глибини дерева;

5) після створення дерев, для кожного елемента рахується середня довжина шляху до гілок дерева;

6) визначаються аномалії за рахунок перевірки значень середньої довжини шляху з певним порогом.

Приклад графічного представлення результатів використання моделі наведено на рисунку 1.7.

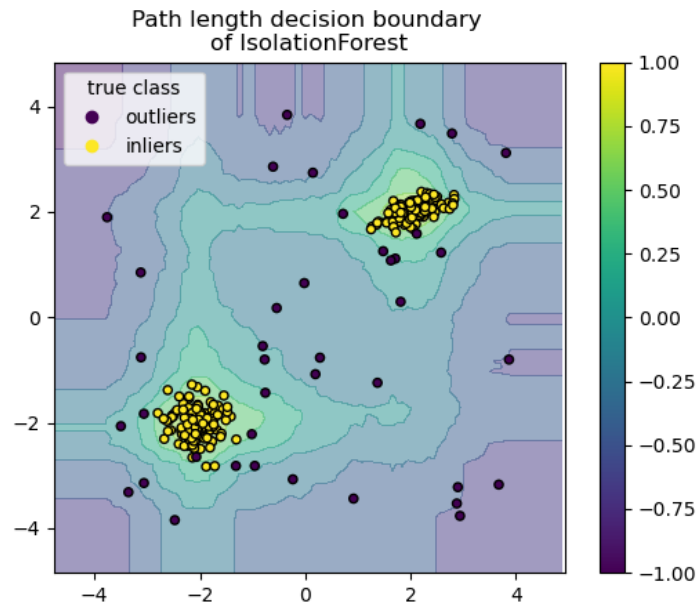


Рисунок 1.7 – Результати роботи методу Isolation Forest (світлими точками позначені нормальні значення, а темними – аномалії) [28]

На рисунку видно, що метод дозволив ізолювати жовті точки в два кластери за рахунок довжини шляху в дереві, а фіолетові точки, в свою чергу, віддалені від цих кластерів і можуть бути аномаліями.

Даний метод може ефективно працювати з великими наборами даних, стійкий до шуму та може досить ефективно знаходити аномальні значення. В той же час, метод може бути схильний до перенавчання за умови великої кількості дерев, що в свою чергу може призвести до зниження точності роботи методу.

Іншим досить розповсюдженим методом є метод k - найближчих сусідів. Даний метод досить поширений в машинному навчанні в першу чергу в задачах класифікації та регресії, проте може використовуватися і для пошуку аномалій [29]. Основний принцип методу полягає в знаходженні відстані від точки до k - найближчих сусідів. Для визначення відстані зазвичай застосовують евклідову відстань, хоча в певних випадках застосовують інші метрики, наприклад, відстань Геммінга. Знаходження евклідової відстані для двох точок можна записати наступною формулою:

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}, \quad (1.4)$$

де x_1 та y_1 - координати першої точки, а x_2 та y_2 – відповідні координати другої точки.

Перевірка чи є певна точка аномальною полягає в порівнянні середнього значення відстаней з певним пороговим значенням. Приклад результату застосування методу наведено на рисунку 1.8.

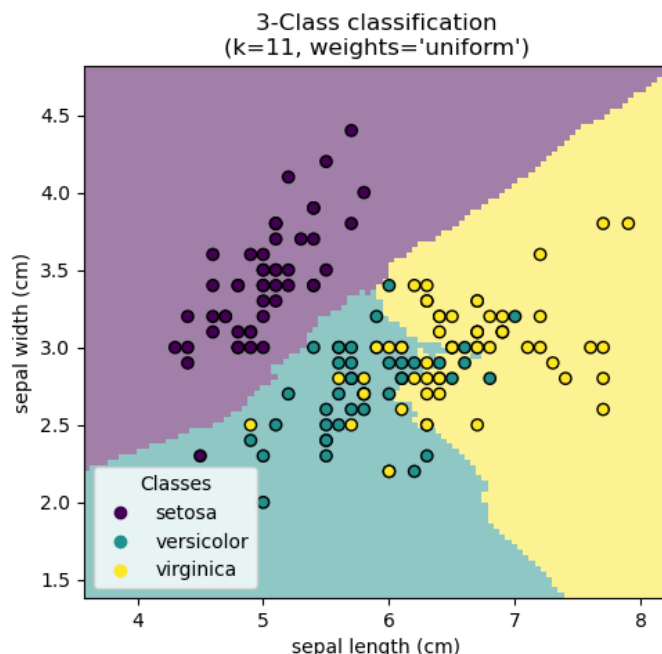


Рисунок 1.8 – Результати роботи методу k -найближчих сусідів (точка певного кольору позначає приналежність до відповідного класу цього самого кольору) [28]

Даний метод може працювати з різними типами даних та при не-однорідному розподілі даних. Головним недоліком напевно є залежність якості роботи від налаштувань, а саме параметру k та порогового значення. Вибір оптимального значення параметру k вимагає гарного знання предметної області та, потенційно, застосування таких технік як перехрестна валідація. Також в літературі рекомендують для визначення параметру знаходити квадратний корінь від розміру усього набору даних.

Широкого застосування для пошуку аномалій мають і методи, які базуються на концепції щільності. Серед таких методів можна виділити коефіцієнт локального відхилення (англійською «Local outlier factor» або просто «LOF») – алгоритм, який базується на попередньому та дозволяє знайти локальну густину точок в наборі [29, 30]. Суть методу полягає в наступному [29, 30]:

1) для кожної точки набору знаходить k – distance – відстань до k -найближчого сусіда (параметр k задається при створенні моделі)

2) отримане значення використовується для знаходження відстані досяжності за формулою:

$$RD_k(A, B) = \max\{k - \text{distance}(B), d(A, B)\}, \quad (1.5)$$

де $d(A, B)$ – відстань від об'єкту A до об'єкту B , а k – distance – відстань об'єкту B до його k -найближчого сусіда;

3) обраховується локальна щільність досяжності об'єкту, яка є оберненою до середньої відстані досяжності відповідного об'єкту до його сусідів. Це позначається за допомогою

$$lrd_k(A) = 1 / \left(\frac{\sum_{B \in N_k(A)} RD_k(A, B)}{|N_k(A)|} \right), \quad (1.6)$$

де $N_k(A)$ – набір k -найближчих сусідів об'єкту A , а $RD_k(A, B)$ – відстань досяжності від об'єкту A до об'єкту B ;

4) знаходиться коефіцієнт локального відхилення (LOF) шляхом відношення середнього значення lrd від k -кількості сусідів точки до lrd цієї точки. Це можна позначити за допомогою наступного виразу:

$$LOF_k(A) = \frac{\sum_{B \in N_k(A)} \frac{lrd_k(B)}{lrd_k(A)}}{|N_k(A)|} = \frac{\sum_{B \in N_k(A)} lrd_k(B)}{|N_k(A)| * lrd_k(A)}, \quad (1.7)$$

де $|N_k(A)|$ – набір k -найближчих сусідів об'єкту A , $lrd_k(B)$ - локальна щільність досяжності об'єкту B , а $lrd_k(A)$ - локальна щільність досяжності об'єкта A ;

5) отримані значення, що дорівнює приблизно 1 вказує на те, що об'єкт можна порівняти зі своїми сусідами, значення нижче 1 вказує на більш щільну область (що вказує на нормальну точку), тоді як значення, значно більші за 1, вказують на аномальні точки.

Даний метод дозволяє знаходити аномалії з врахуванням локальних властивостей набору даних і при цьому може працювати з різними формами розподілу даних. Загалом він дає кращі результати за глобальний підхід до пошуку, проте немає загального порогового значення LOF , тому вибір допустимого відхилення не найпростіша задача. Приклад результатів застосування методу наведено на рисунку 1.9.

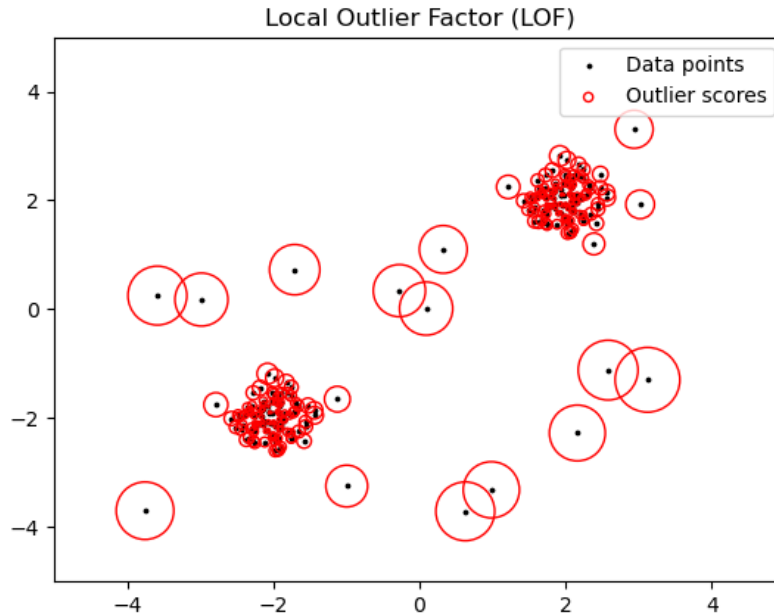


Рисунок 1.9 – Результати роботи методу LOF [28]

Досить цікавим методом є Seasonal Hybrid ESD, який дозволяє знаходити аномалії в часових рядах з сезонністю [31]. Тест Грабса, також відомий як екстремальний студентизований тест відхилень (ESD) є тестом, який використовується для виявлення аномалій у одномірному наборі даних, підпорядкованому нормальному розподілу. Тест Грабса можна позначити:

$$G = \frac{\max_{i=1 \dots N} |Y_i - \bar{Y}|}{s}, \quad (1.8)$$

де $\bar{Y}$ – це середнє значення вибірки, а s – стандартне відхилення.

Проблема тесту Грабса в тому, що він передбачає нормальний розподіл, проте дані далеко не завжди розподілені по такому закону. Метод Seasonal Hybrid ESD пропонує використовувати медіану та середнє абсолютне відхилення замість середнього та стандартного відхилення, оскільки вони дуже чутливі до великих та чисельних аномалій. Він комбінує сезонну адаптацію з експоненційно зваженою середньою (Exponential Smoothing) і використовує Extreme Studentized Deviate (ESD) для виявлення відхилень від очікуваного розподілу даних. Переваги моделі такі:

- врахування сезонності: враховує сезонні зміни в часових рядів, що дозволяє ефективно виявляти аномалії, які повторюються у певних періодах часу;

- гнучкість та налаштування: метод може бути налаштований залежно від потреб користувача (період сезонності та порогові значення);

- підтримка широкого спектра даних: може бути застосований до різних типів даних, включаючи числові часові ряди, а також категоріальні та багатовимірні дані.

Серед недоліків можна виділити:

- обмежена робота з незвичайними аномаліями: метод краще працює з аномаліями, які досить схожі на звичайні сезонні зміни;

- вимоги до часу обчислень: алгоритм має не найбільшу швидкість, особливо для великих обсягів даних або довгих часових рядів.

1.5.2. Огляд засобів на Python для автоматизації пошуку аномалій в часових рядах

Як було зазначено вище, існує великий перелік метод та підходів для пошуку аномалій в часових рядах. В переважній більшості випадків науковці не проводять необхідні розрахунки вручну, а застосовують автоматизовані шляхи. Мова програмування Python, беззаперечно, є найбільш популярною мовою для вирішення задач системного аналізу. Однією з основних причин цієї популярності є наявність широкого набору бібліотек та інструментів, які, в свою чергу, програмно реалізують усі раніше згадані методи. Виходячи з цього, варто розглянути декілька подібних бібліотек та проаналізувати їх ефективність.

Для огляду буде використано дані моніторингу стану атмосферного повітря за даними мережі ECoCity. Дані для дослідження отримані за допомогою сервісу «Кабінет дослідника» [19], який був детально описаний в пункті 1.1.4. В цьому конкретному випадку було обрано показник «PM10» (пил, розміром 10 мкм і менше), отриманих з однієї зі станцій у Вінницькій області. На рисунку 1.10 зображено графік даного часового ряду.

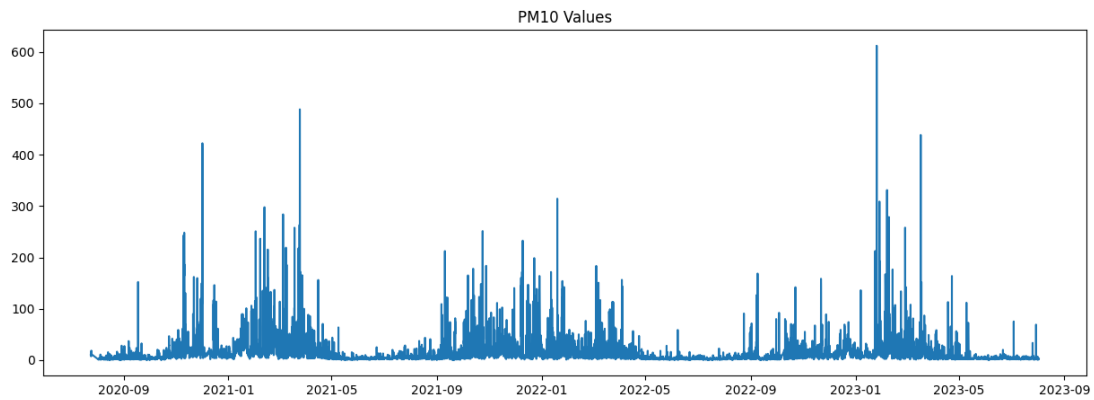


Рисунок 1.10 – Вхідні дані для передбачення аномалій за показником PM10 у Вінницькій області [32]

Проаналізуємо можливості Python-бібліотеки `sesd`. Дана бібліотека реалізує метод Seasonal Hybrid ESD, який був детальніше описаний в попередньому пункті. Приклад коду для виявлення аномалій, використовуючи бібліотеку `sesd` (Seasonal Hybrid ESD) наведено на рисунку 1.11, а результат його роботи – на рисунку 1.12.

```
anomaly_df = df['y']
outliers_indices = sesd.seasonal_esd(anomaly_df,
                                     hybrid=True,
                                     max_anomalies=2500,
                                     alpha = 3)
anomalies = df.loc[outliers_indices]
```

Рисунок 1.11 – Приклад використання бібліотеки `sesd` [32]

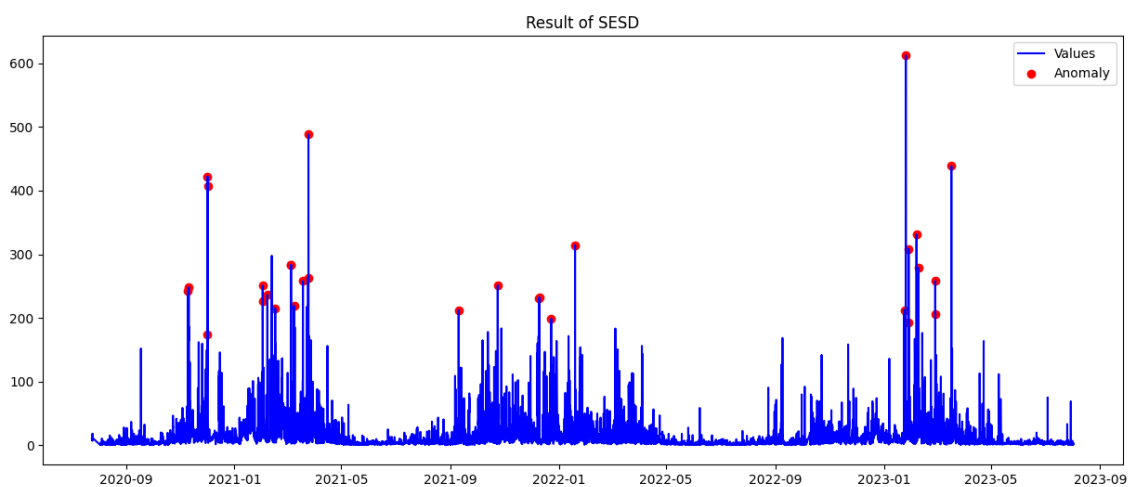


Рисунок 1.12 – Результат роботи бібліотеки `sesd` [32]

Наступною розглянемо бібліотеку scikit-learn [28]. Дана бібліотека містить велику кількість методів і моделей машинного навчання як для задачі регресії і кластеризації, так і для пошуку аномалій. В даному огляді обмежимося модулем Isolation Forest, який реалізовує однойменний метод пошуку аномалій, що був детально описаний в попередньому пункті. Приклад коду застосування даного методу наведено на рисунку 1.13, а результат роботи – на рисунку 1.14.

```
model = IsolationForest(contamination=0.05)
model.fit(df['y'].values.reshape(-1, 1))
y_pred = model.predict(df['y'].values.reshape(-1, 1))
```

Рисунок 1.13 – Приклад використання модуля Isolation Forest Python- бібліотеки scikit-learn [32]

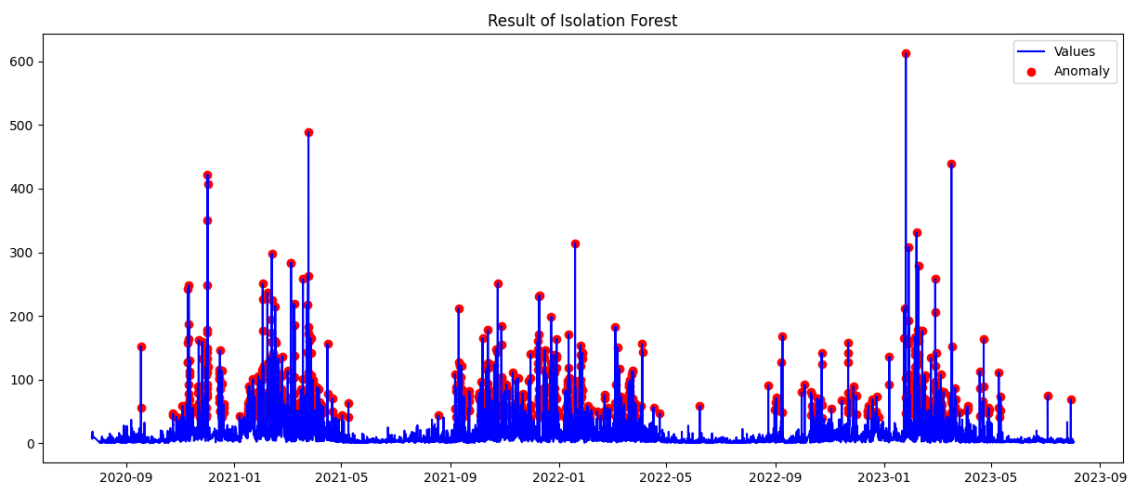


Рисунок 1.14 – Результат роботи Isolation Forest [32]

Розглянемо бібліотеку statsmodels. У statsmodels немає вбудованих функцій для прямого виявлення та фільтрації аномалій у часових рядах. Однак, можна використовувати інші функції та методи з бібліотеки для реалізації такого аналізу. Один з підходів – це аналіз залишків. Можна використовувати статистичні моделі, такі як ARIMA або ETS, для прогнозування часового ряду. Порівнюючи фактичні значення з прогнозованими, можна виявити аномалії, коли значення значно відрізняються від прогнозу. Інший підхід – це використання статистичних методів. Це включає встановлення порогових значень, використання стандартного відхилення, Z-перетворення та інших методів для виявлення значень, які перевищують певний поріг.

Переваги використання statsmodels включають наявність статистичних моделей та гнучкість у виборі підходу для виявлення аномалій. Однак, для ефективного використання бібліотеки потрібні знання статистики та часових рядів. Спробуємо знайти аномалії, використовуючи розклад часових рядів, як це показано на рисунку 1.15 з результатами, зображеними на рисунку 1.16.

```
period = 24
result = seasonal_decompose(df_anomaly['y'], model='additive', period=period)
residual = result.resid
threshold = 200.0
anomalies = df_anomaly[residual.abs() > threshold]
df_filtered = df_anomaly.drop(anomalies.index)
```

Рисунок 1.15 – Приклад використання бібліотеки statsmodels [32]

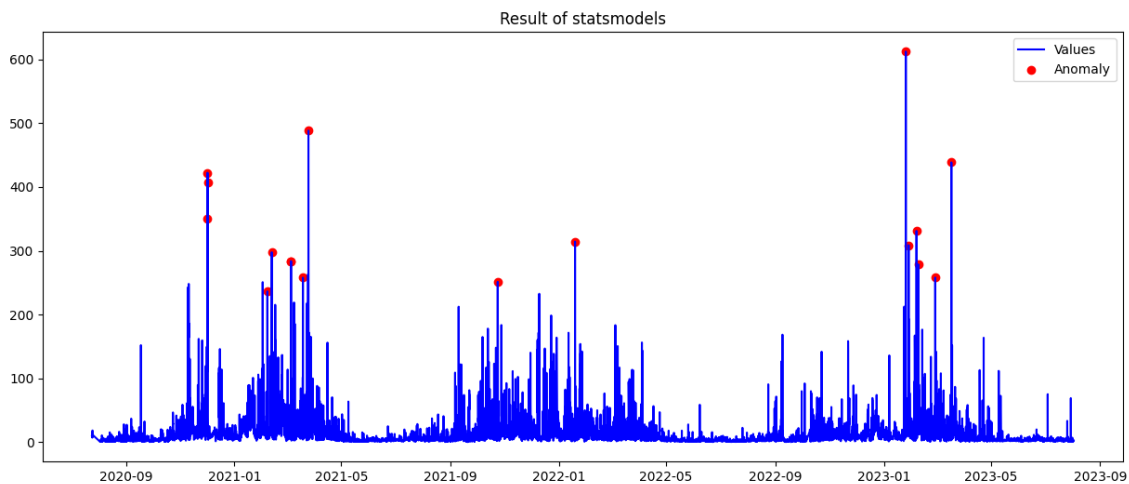


Рисунок 1.16 – Результати, отримані за допомогою бібліотеки statsmodels [32]

Як показав аналіз, мова програмування Python дійсно має великий перелік бібліотек, серед яких ті, що були безпосередньо описані в даному пункті. Відповідно, це дозволяє доволі швидко проводити дослідження, пов'язані з пошуком аномалій відомими, оскільки усуває потребу в програмній реалізації відповідного методу.

1.6. Систематизація методів розкриття невизначеності для різних видів декомпозиції та постановка задач дослідження

У підрозділі 1.2 здійснено декомпозицію складної системи за різними критеріями та виділено види невизначеності для їх елементів. У наступних пунктах і підрозділах наведено огляд та аналіз підходів, методів і технологій для розкриття більшості з цих видів невизначеності.

Систематизуємо їх у таблиці 1.3 та визначимо які варіанти розкриття невизначеності ще потребують удосконалення та які методи ще необхідно розробити чи удосконалити.

Таблиця 1.3 – Методи розкриття невизначеності для різних видів декомпозиції системи управління станом атмосферного повітря регіону

№ п/п	Вид декомпозиції	Елемент декомпозиції	Вид невизначеності	Сутність невизначеності	Підхід чи метод розкриття невизначеності
1.	D_{DPSIR}	D_S	U_{T2}	невизначеність щодо критеріїв якості атмосферного повітря в цілому	Використання індексу якості повітря
2.	D_{DPSIR}	D_D	U_{I1}	неповна чи неточна інформація про місця найбільшої щільності розташування підприємств-забруднювачів	Інформаційний запит
3.	D_{DPSIR}	D_D	U_{I2}	неповна чи неточна інформація про найбільші обсяги гранично допустимих значень забруднювальних речовин у викидах	Інформаційний запит
4.	D_{DPSIR}	D_P	U_{I3}	неповна чи неточна інформація про місця розташування джерел викидів та про фактичні викиди	Інформаційний запит
5.	D_{DPSIR}	D_S	U_{I4}	неточні дані про значення показників якості атмосферного повітря	Використання даних систем моніторингу
6.	D_{DPSIR}	D_R	U_{I5}	неповна чи неточна інформація про вже прийняті рішення	Інформаційний запит
7.	D_{DPSIR}	D_R	U_{T4}	невизначеність вибору рішень	Інформаційний запит
8.	D_{DPSIR}	D_S	U_{M1}	невизначеність щодо варіантів моделей складної системи	Огляд інтелектуальних моделей прогнозування

Продовження таблиці 1.3

№ п/п	Вид декомпозиції	Елемент декомпозиції	Вид невизначеності	Сутність невизначеності	Підхід чи метод розкриття невизначеності
9.	D_{DPSIR}	D_P	U_{S1}	відсутність достовірної інформації про перелік усіх забруднювачів, через те, що якесь підприємство здійснює викид, але не подає про це звітність, або фальсифікує інформацію про реальні викиди	Метод пошуку можливих джерел забруднення
10.	D_{DPSIR}	D_P	U_{S1}	відсутність достовірної інформації про перелік усіх забруднювачів, через те, що якесь підприємство здійснює викид, але не подає про це звітність, або фальсифікує інформацію про реальні викиди	Метод пошуку можливих джерел забруднення
11.	D_{DPSIR}	D_S	U_{T3}	має місце невизначеність щодо вибору варіантів цього порогу	Використання директив ЄС
12.	$D_T(\Delta T)$		U_{T5}	невизначеність щодо вибору варіантів цього порогу	Використання директив ЄС
13.	D_{DPSIR}	D_S	U_{T1}	невизначеність вибору місць розташування існуючих чи нових постів моніторингу	Використання методу аналізу ієрархій
14.	$D_{SEC}(\Theta)$		U_{P2}	невизначеність щодо параметрів моделей, які описують вплив сезонної складової: період, вікно впливу, коефіцієнт регуляризації або сила впливу	Метод ідентифікації параметрів сезонності та аномалій
15.	D_{DPSIR}	D_S	U_{P1}	невизначеність щодо значень параметрів моделей	Метод ідентифікації параметрів сезонності та аномалій

Продовження таблиці 1.3

№ № п/п	Вид деко- позиції	Елемент деком- позиції	Вид невизна- ченості	Сутність невизначеності	Підхід чи ме- тод розкриття невизначеності
16.	$D_{SEC}(\Theta)$		U_{T6}	вибір яка саме сезонна складова або їх сукупність мають місце	Метод ідентифікації параметрів сезонності та аномалій
17.	$D_T(\Delta T)$		U_{S2}	у різні інтервали часу можуть бути різні аномалії чи пікові значення, які важко виявляти	Метод ідентифікації локальних у часі аномалій
18.	D_{DPSIR}	D_I	U_{ST1}	неповне знання про елементи системи або неточність уявлень про зв'язки між ними	Байєсівське моделювання
19.	D_G	D_{G-O}	U_{ST2}	неповне знання про взаємозв'язок між різними показниками стану атмосферного повітря на різних постах спостережень	Байєсівське моделювання
20.	D_{DPSIR}	D_I	U_{SC1}	відсутність чи неточність прогнозів динаміки системи для різних сценаріїв розвитку	Застосування алгоритму системного аналізу стану повітря

Виходячи зі списку, наведеного у пункті 1.3.4 та таблиці 1.3, можна зробити висновок, що для розкриття деяких видів невизначеності достатньо використати інформації з відкритих джерел, таких як інформація мереж моніторингу, звіти відповідних статистичних органів, тощо. Проте, для переважної більшості елементів необхідно розробити або удосконалити наступні методи:

1. Запропонувати метод пошуку можливих джерел підвищеного забруднення на основі аналізу аномалій, який повинен надавати можливість локалізувати у часі та просторі джерела сталого негативного ефекту на рівень забруднення, аналізуючи наявні дані моніторингу стану атмосферного повітря.

2. Розробити метод формування системи критеріїв та прийняття рішення щодо оптимального розміщення станцій мережі моніторингу, застосовуючи метод аналізу ієрархій та врахуванням аномалій показників стану атмосферного повітря.

3. Запропонувати метод ідентифікації параметрів гармонік та аномалій періодичного часового ряду на основі адаптивної декомпозиції.

Метод повинен дозволити знайти та ефективно комбінувати декілька варіантів сезонності при роботі з інтелектуальними моделями прогнозування стану атмосферного повітря.

4. Запропонувати метод ідентифікації локальних у часі аномалій за результатами декомпозиції ряду, що має дозволити ідентифікувати окремі інтервали часового ряду та виконати на ньому пошук аномальних значень.

5. Удосконалили метод аналіз взаємовпливу показників стану атмосферного повітря різних регіонів на основі байєсівського моделювання.

Розв'язанню цих задач присвячені наступні розділи цієї роботи.

1.7. Висновки до розділу 1

Наведено опис стану атмосферного повітря в Україні, директиви та методи контролю за його станом. Проаналізовано основні показники якості атмосферного повітря. Здійснено огляд європейського індексу якості повітря та порівняно з українським аналогом. Описано мережі державного та громадського моніторингу, яку функціонують на території України.

Описано методологію системного аналізу у складних системах та за цією методологією охарактеризовано об'єкт дослідження. Проведено декомпозицію системи за різними критеріями та класифіковано види невизначеності, які необхідно розкрити в рамках роботи.

Проведено порівняльний аналіз методів та інтелектуальних моделей прогнозування часових рядів стану атмосферного повітря з урахування методів пошуку аномалій в цих рядах.

Описано основні проблеми системного аналізу та прогнозування стану атмосферного повітря. Поставлено задачі, які необхідно розв'язати. Безпосередньо їх розв'язанню присвячені наступні розділи монографії.

Визначено та систематизовано перелік задач, які варто розв'язати для розкриття невизначеності для різних видів декомпозиції.

РОЗДІЛ 2

СИСТЕМНИЙ АНАЛІЗ СТАНУ АТМОСФЕРНОГО ПОВІТРЯ З УРАХУВАННЯМ ВПЛИВУ АНОМАЛІЙ

2.1. Математична постановка задачі системного аналізу

Проведений у розділі 1 аналіз дозволяє уточнити математичну постановку задачі управління станом атмосферного повітря з використанням методології системного аналізу.

В контексті 5-ти етапів системного аналізу, як було вказано в підрозділі 1.2, етап 5 щодо оптимізації складної системи та підтримки прийняття рішень у сфері управління станом атмосферного повітря регіоном вже є добре вивченим, тому потребують актуалізації тільки етапи 1-4. Етап 4 полягає у проведенні інтелектуального аналізу даних по визначеній на етапі 3 оптимальній моделі: проведення прогнозування, програвання сценаріїв, пошук залежності між ознаками датасету, вивчення зв'язку між різними станціями моніторингу стану атмосферного повітря. Отже, ключовим є етап 3, де необхідно побудувати оптимальну модель. Задача прогнозування має регресійний характер, оскільки модель будується для прогнозування дробових значень цільових ознак Y , якими є показники стану атмосферного повітря або інтегральний індекс якості атмосферного повітря (AQI) у заданому регіоні. Змінними стану X , по яких обчислюються ці вихідні змінні Y , є значення цих же та інших показників у попередні періоди часу.

Входами U є антропогенні джерела забруднень, яке надходить з інших регіонів або з джерел забруднення у цьому регіоні. Тільки цими факторами можна ефективно керувати через дозволи на гранично допустимі викиди та систему штрафів, через їх порушення. В системі мають місце обмеження: джерела забруднення не можуть викидати менше нуля. В ідеалі, вони повинні здійснювати викиди в межах між мінімально зручними для функціонування джерела $U_{\min}$ і максимально допустимими $U_{\max}$ значеннями, які встановлюються гранично допустимими дозволами на викиди (ГДВ): $U_{\min} \leq U \leq U_{\max}$.

Збуренням чи завадами ξ є метеорологічні та інші фактори, які впливають на стан атмосферного повітря, окрім джерел забруднень антропогенного характеру.

Для оцінювання оптимальності моделі J пропонується використовувати типові для регресійних задач метрики [5]: коефіцієнт детермінації R^2 (на Python – `r2_score`, її треба максимізувати, ідеальне значення – 1,0), відносну похибку (на Python є її аналог з усередненням на усьому

діапазоні даних – MAPE), корінь із середньоквадратичної похибки (RMSE) тощо. MAPE та RMSE треба мінімізувати, ідеальне значення – 0, тому краще щодо метрик писати, що їх треба «оптимізувати». Щодо коефіцієнту детермінації, то, зазвичай, критерієм є $R^2 \geq 0.7$, тобто моделі, які забезпечують менше значення, вважаються недостатньо адекватними [5].

Таким чином, задачу оптимізації формулюємо таким чином: оптимізувати метрику(и) J для прогнозів показників стану атмосферного повітря (по інтегральному AQI в цілому та по окремих показниках зокрема) для заданого регіону у валідаційному наборі Y , сформованому за даними державного або громадського моніторингу атмосферного повітря цього регіону за даними тренувального набору, який складається з цього та інших показників у попередні періоди часу X , даних про антропогенні джерела забруднення U , даних про збурення чи завади ξ . За побудованою у такий спосіб оптимальною моделлю можна програвати сценарії з пошуку оптимальних значень параметрів антропогенних джерел забруднення U , з урахуванням обмежень $U_{\min} \leq U \leq U_{\max}$, за результатами яких формуються Ω_U рекомендації щодо підтримки прийняття рішень у системі управління атмосферним повітрям регіону. Задача доповнюється необхідністю розкрити основні види невізначеності у ній та провести інтелектуальний аналіз з використанням побудованих за різними критеріями оптимальних моделей і вивчити важливі закономірності щодо взаємозв'язків між усіма видами ознак та даними, що дозволить сформулювати додаткові рекомендації Ω_Y .

2.2. Метод пошуку можливих джерел підвищеного забруднення на основі аналізу аномалій

Як було зазначено в першому розділі, існує проблема з відсутністю достовірної інформації про перелік усіх джерел забруднення, через те, що підприємства, які здійснюють викид забруднювальних речовин, або фальсифікують інформацію про реальні викиди, або взагалі не звітують щодо цього (невизначеність U_{S1}). Це ускладнює процес боротьби з підвищенням рівнів забруднення, оскільки немає можливості визначити що є безпосередньою причиною росту рівня забруднення та в який момент розпочався цей вплив.

Для вирішення цієї задачі, необхідно розробити метод, що дозволить за наявними даними моніторингу локалізувати у просторі та часі джерела забруднення, а в подальшому використати цю інформацію для порівняння з наявними дозволами на викиди забруднювальних речовин і, за потреби, вжити обґрунтовані управлінські рішення.

Пропонується метод, що базується на використанні моделі Prophet, яку було описано у підрозділі 1.4. Дані моніторингу атмосферного повітря, зазвичай, мають різну періодичність та містять пропуски в значеннях спостереження. Модель Prophet досить ефективно працює за таких умов, тому її використання є доцільним.

Загалом, алгоритм розв'язання задачі можна представити наступними кроками [33]:

1. Обираються дані моніторингу якості повітря, які будуть використовуватися як тренувальні (для налаштування параметрів моделі Prophet) та як контрольні значення (для перевірки ефективності роботи методу).

2. Вибираються параметри моделі Prophet (варіанти сезонності, порядок коефіцієнтів рядів Фур'є для їх опису) та ін.

3. Встановлюються точки зміни тренду (англійською «changepoints» або «CP») на 1 число кожного місяця і будується лінійна модель Prophet між цими CP для показника якості повітря на різних постах спостережень окремо, таким чином, щоб вона дозволяла виявляти початок забруднюючого впливу джерел із тренувальної вибірки. Вплив визначається за рахунок аналізу нахилу тренду для кожного такого випадку (між кожними CP для кожного показника на кожному пості) та знаходяться серед них найбільші прирости, тобто найбільші нахили, але сходинкового типу, тобто, коли мало місце збільшення навантаження по заданому показнику, виявлене на певному пості, а потім воно не зменшувалось, а тільки наростало аж по останній день, тобто ті випадки коли, наприклад, підприємство розпочало роботу та, відповідно, забруднюючий вплив, так і продовжує це робити. Результатом аналізу є приріст забруднення у дату, з якої все почалось, решта приростів, де ця умова не виконується, прирівнюються до нуля.

4. Виявлені дати та забруднюючий вплив порівнюється із контрольними даними, відібраними у пункті 1, за певною метрикою. Змінюємо структуру і параметри моделі та повторюємо пункти 2 та 3, поки похибка не стане прийнятною, або робимо повний перебір допустимих варіантів моделі і вибираємо модель з найменшою похибкою.

В результаті отримуються місця та дати початку забруднюючого впливу. Цю інформацію можна порівняти, з відомими даними щодо викидів забруднювальних речовин та відповідними дозволами на викиди цих речовин. Потім інформація про локалізовані у часі та просторі незареєстровані джерела забруднення передаються відповідним контролюючим органам для перевірки.

Для підвищення точності роботи технології варто на етапі 2 використати параметр «holidays». Даний параметр дозволяє вказати моделі

перелік дат відомих аномальних впливів, таких як природні або антропогенні катастрофи, які можуть суттєво впливати на рівень забруднення у відповідний момент часу та тим самим погіршувати точність роботи моделі. Важливим є також те, що вплив таких забруднень може бути довготривалим, тобто у наборі «holidays» слід правильно вказувати «вікно» тривалості такого впливу.

Важливо зазначити, що цей метод не розрізняє причини забруднюючих впливів (антропогенні або природні), а тільки співставлення з наявними даними та подальше вивчення результатів технології дозволяє ідентифікувати їх природний чи антропогенний характер.

Метод є дуже чутливим і дозволяє зафіксувати, навіть, незначне підвищення тренду, але таке, що значно відрізняється від тренду у попередні моменти часу.

Ключовим у цьому алгоритмі, окрім застосування моделі Prophet та правильного налаштування її параметрів, є алгоритм виявлення імпульсів, які характеризують початок появи сталого забруднення, тобто такого, що має місце тривалий час після появи, та актуального, тобто такого, що має місце і по цей час, тому зупинимось на ньому більш детально.

Модель Prophet повертає початковий тренд k_0 ряду до першої точки зміни d_0 і m приростів $\Delta k_i, i = \overline{1, m}$ тренду між наступними m точками аж до кінця тренувального інтервалу даних. За необхідності, потім по цій моделі можуть передбачатись дані на наступні періоди. При цьому, дати $d_i, i = \overline{1, m}$ розташування цих точок вже не зазнають змін.

Головним є знаходження моментів приростів, тому початковий тренд значення не має і слід аналізувати тільки прирости тренду відносно нього. Результатом роботи моделі Prophet є

$$\hat{y}_i = (k + \Delta k_i) d_i, i = \overline{1, m}, \quad (2.1)$$

де $\hat{y}_i$ – стандартизоване значення тренду в i -й точці зміни d_i , тобто масштабоване між мінімальним і максимальним значеннями та поділене на середньоквадратичне відхилення – це стандартне позначення пакету програм Prophet, яке у результатах так і називається: “yhat” (рис. 2.1).

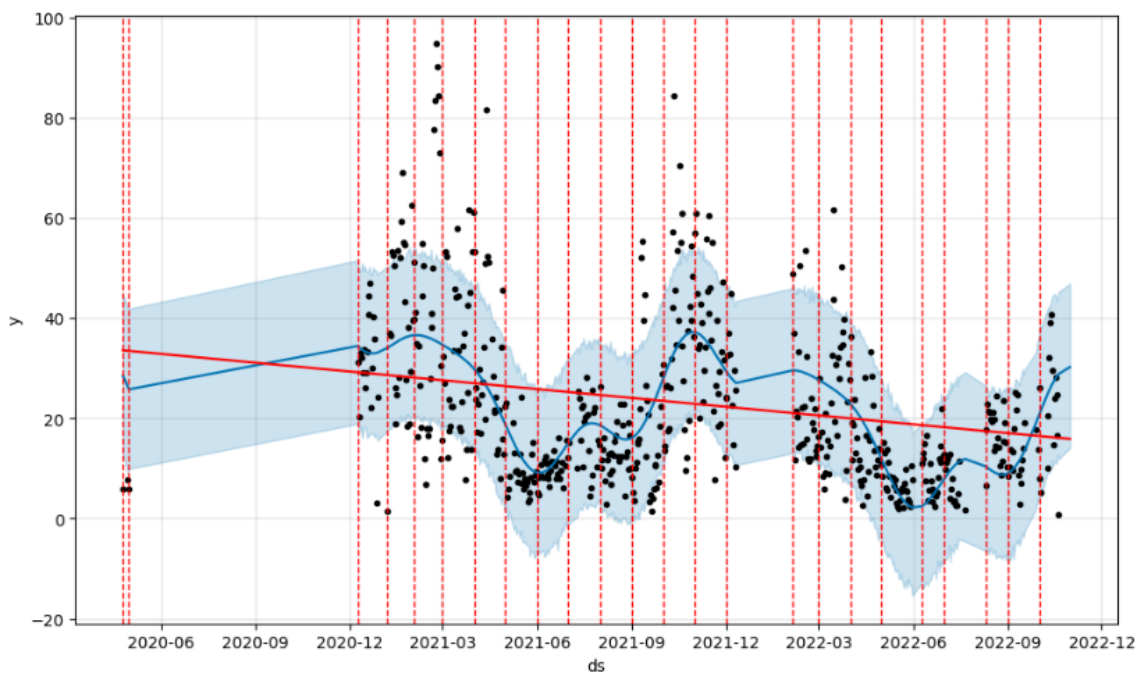


Рисунок 2.1 – Значення частинок PM10 на посту моніторингу у місті Вінниця за 2020-2022 рр. та результат роботи моделі з зоною невизначеності та трендом [33]

На рис. 2.2 показано прирости Δk_i тренду між точками зміни d_i .

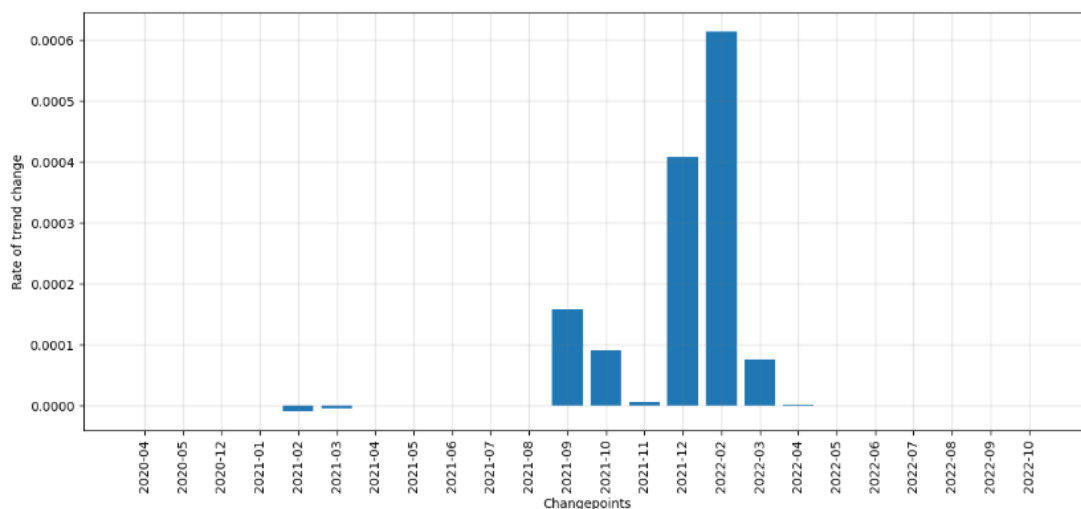


Рисунок 2.2 – Значення приростів Δk_i тренду показника PM10 на посту в місті Вінниця за 2020-2022 рр. між точками зміни d_i [33]

Необхідно сформулювати умову для виявлення ділянок, де прирости трендів є суттєвими (наприклад, більшими, ніж на 5% від максимального значення). Як відомо, для виявлення імпульсів зазвичай використовується інтегрування. Здійснимо інтегрування, тобто знайдемо S суми приростів тренду з накопиченням:

$$S = \sum_{i=1}^m \Delta k_i . \quad (2.2)$$

На рисунку 2.3 наведено результат оброблення даних рисунку 2.2 за формулою (2.2).

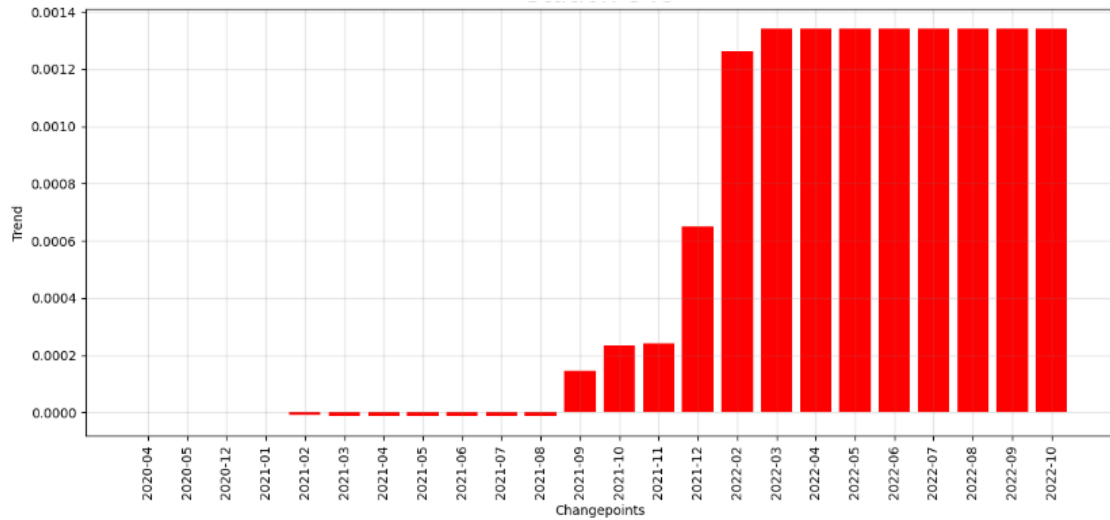


Рисунок 2.3 – Значення сум приростів тренду з накопиченням показника РМ10 на посту в місті Вінниця за 2020-2022 рр. між точками зміни d_i [33]

Але аналіз показує, що ряд прирості Δk_i тренду може бути знаковим. У той же час, важливо знайти саме такі імпульси, вплив яких є і досі актуальним, а не епізодичним. Тобто треба знайти ситуації, коли після такого імпульсу концентрація далі весь час тільки наростала. Значення S можуть змінюватись синусоїдально і по них важко буде однозначно виявити місця постійного наростання концентрації. Для цього пропонуємо здійснити ще одне інтегрування (рис. 2.12):

$$J = \sum_{j=1}^m S_i = \sum_{q=j}^m \sum_{j=1}^m \Delta k_{qj} , \quad (2.3)$$

а тоді слід лише знайти такі імпульси на графіку на рис. 2.4, після яких функція J далі лише монотонно наростала.

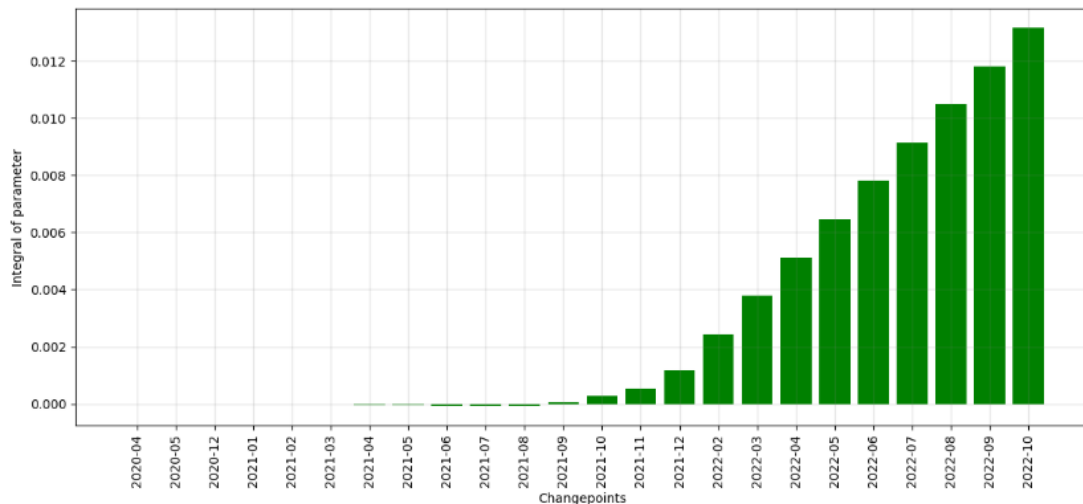


Рисунок 2.4 – Значення сум J приростів тренду з накопиченням показника PM10 на посту в місті Вінниця за 2020-2022 рр. між точками зміни d_i [33]

Таким чином алгоритм методу пошуку можливих джерел підвищеного забруднення на основі аналізу аномалій визначається так:

1. Вибрати показники моніторингу якості повітря, для яких часто має місце перевищення гранично допустимих значень у контрольній точці та по яких немає впевненості щодо джерел їх походження.

2. Сформувати та очистити датасети від помилкових та пропущених даних.

3. Сформувати набір даних з датами відомих зафіксованих викидів забруднювальних речовин (наприклад, антропогенні або природні катастрофи), в яких можливі перевищення показника, що досліджуються.

4. Виділити з кожного датасету тренувальні дані (75-80%), на яких будуть ідентифікуватись параметри моделі, і тестові дані, на яких буде вибиратись оптимальна модель з числа ідентифікованих на тренувальних даних.

5. Задати початкове наближення точок зміни тренду «changepoints», встановивши їх на першу дату кожного місяця з числа тренувальних даних.

6. Здійснити налаштування моделі на основі технології Prophet з лінійним трендом, з врахування даних типу «holidays» та врахуванням сезонності, яка мінімізує похибку на тестових даних. Побудувати різні моделі, варіюючи такі параметри:

- Адитивний чи мультиплікативний варіант врахування тренду і сезонних компонент моделі;

- Різне врахування сезонності – тільки річна, яка є обов'язковою, чи ще й по порах року (період 365,25/4);

- Різні значення коефіцієнтів рядів Фур'є, якими апроксимуються сезонні компоненти ряду;

– Різна кількість точок зміни тренду.

7. Вибрати оптимальну за структурою і параметрами модель, ідентифіковану на етапі 6, яка забезпечить мінімальну похибку на тестових даних, відібраних на етапі 4.

8. Для кожного i -го показника для кожної ділянки між j -ми постами спостережень і для усіх дат між k -ми точками «changepoints» показника якості повітря x_{ijk} , масштабованого по цьому показнику, визначити бальну y_{ijk} оцінку приросту тренду x цього показника, яка більшим значенням дає значно більшу вагу, наприклад, шляхом взяття квадрату від їх значень:

$$y_{ijk} = \left(\frac{x_{ijk} - \min_i x_{ijk}}{\max_i x_{ijk} - \min_i x_{ijk}} \right)^2, \quad (2.4)$$

або більш простий варіант для випадку, коли мінімальне значення береться таким, що дорівнює нулю:

$$y_{ijk} = \left(\frac{x_{ijk}}{\max_i x_{ijk}} \right)^2. \quad (2.5)$$

9. Для кожної k -ї дати кожної ділянки між j -ми постами спостережень слід додати усі ці бальні оцінки та ранжувати їх за зменшенням

$$J_{jk} = \sum_i y_{ijk}. \quad (2.6)$$

10. Проаналізувати місця та дати з найбільшими значеннями J_{jk} з урахуванням даних звітності, дозволами на викиди забруднювальних речовин, результатами перевірок, даними супутникових спостережень, даними громадського екологічного моніторингу та іншою інформацією, здійснити пошук, коли у визначені дати на визначених ділянках могли з'явитись джерела забруднення, та рекомендувати відповідним органам провести більш ретельну перевірку виявлених фактів.

Блок-схема алгоритму даного методу наведена на рисунку 2.5.

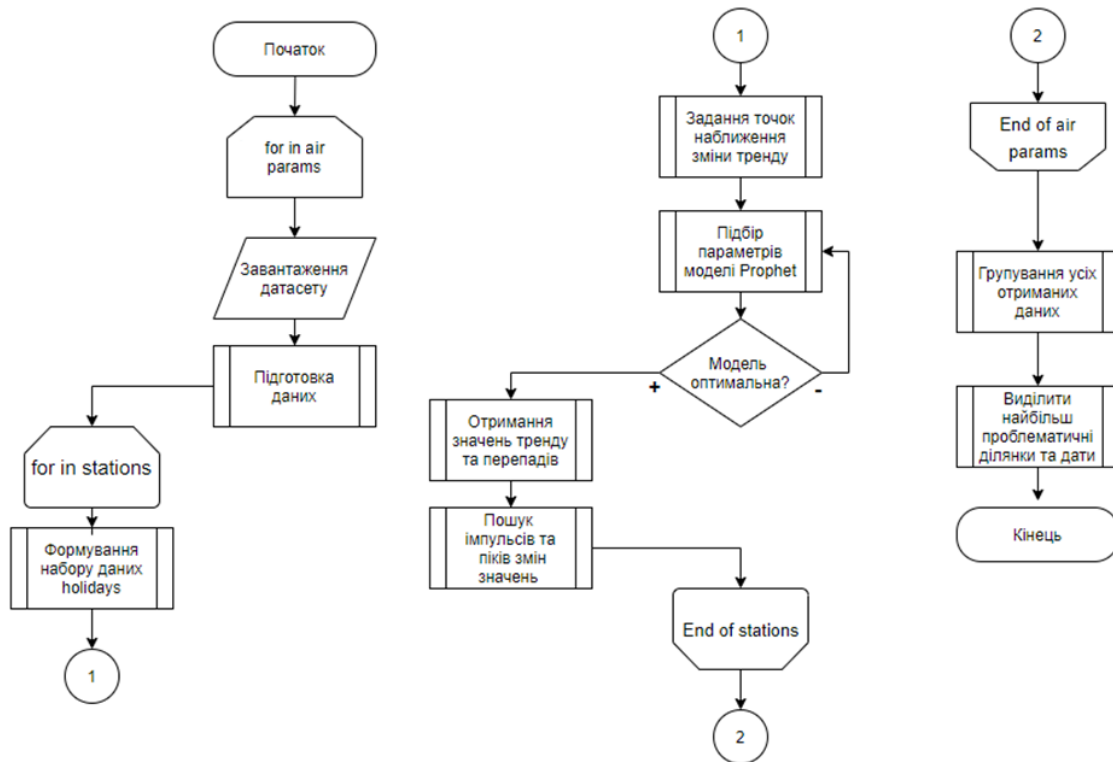


Рисунок 2.5 – Блок-схема алгоритму методу пошуку можливих джерел забруднення [33, 34]

Перевіримо цей метод для даних моніторингу якості атмосферного повітря. Для цього використаємо дані щодо стану атмосферного повітря від мережі громадського моніторингу EcoCity, отримані з декількох станцій моніторингу атмосферного повітря, розташованих у Вінницькій області (рис. 2.6).

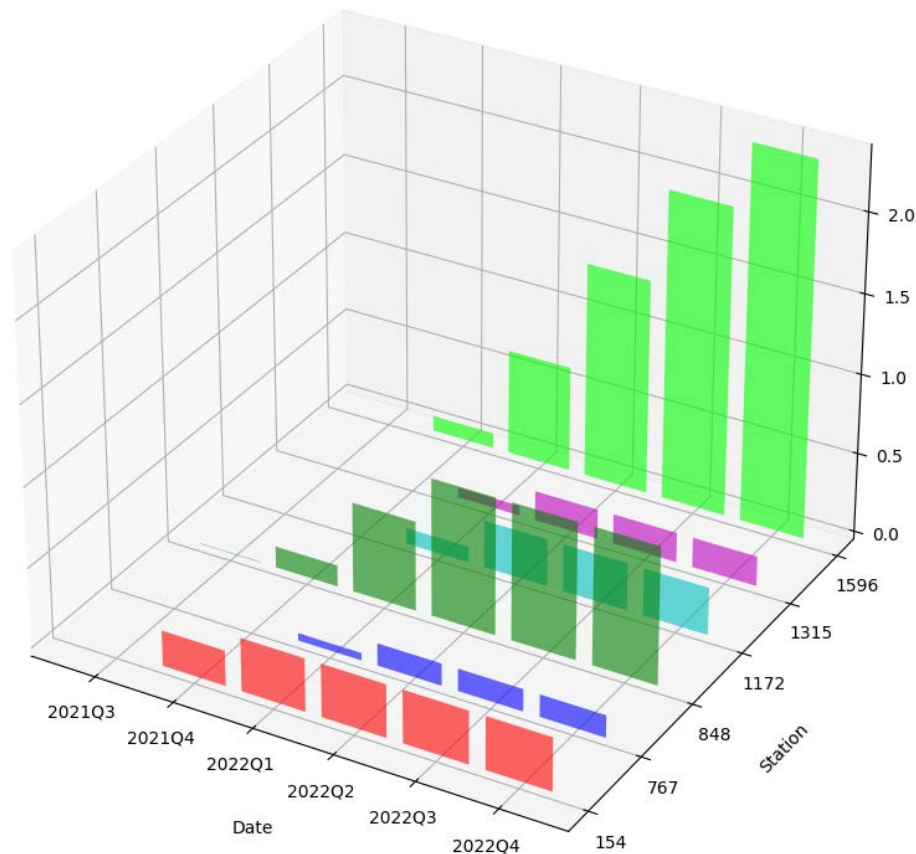


Рисунок 2.6 – Результат пошуку аномальних впливів у часі у Вінницькій області за середньоквартальними даними 2020-2022 рр. [34]

З рисунку 2.6 видно, що на станції № 1596 (м. Хмільник), починаючи з 4-го кварталу 2021 року, розпочався сталий вплив забруднення, що продовжувався місяцями. Отримані результати можуть бути передані в Держекоінспекцію у Вінницькій області для порівняння з інформацією про дозволи на викиди та прийняття відповідних рішень.

Отже, метод пошуку можливих джерел забруднення на основі аналізу аномалій є ефективним інструментом, який можна успішно застосовувати для вирішення поставленої задачі і який дозволяє розкрити невизначеність U_{S1} .

2.3. Метод формування системи критеріїв для оптимального розміщення станцій мережі моніторингу

В підрозділі 1.6 було визначено, що існує проблема з вибором місць розташування існуючих чи нових постів моніторингу систем атмосферного повітря (невизначеність U_{T1}), оскільки існуючі системи можуть збирати недостатню кількість даних для проведення ефективного аналізу стану атмосферного повітря. Відповідно, потрібно покращувати та оптимізувати цю систему, але цей процес є довгим і досить

дорогим, тому необхідно сформулювати та описати метод, який дозволить визначити пріоритетні місця для встановлення станцій, що в далекій перспективі дозволить системним аналітикам більш точно аналізувати та прогнозувати стан атмосферного повітря в межах розташування відповідної мережі моніторингу.

2.3.1. Визначення екстремальних значень показників про стан атмосферного повітря

Як було зазначено вище, відповідно до чинного законодавства [12], аналіз даних спостережень стану атмосферного повітря розпочинається з порівняння показників з пороговими значеннями.

При порівняння показників з пороговими значеннями беруть до уваги такі умови:

- Середнє значення на рік (для показників тверді частки (PM10/PM2.5), важкі метали, Бенз(а)пірен та ін.) X_{AVG} ;

- Максимальне середньодобове восьмигодинне значення (для показника озон) X_{MAX} ;

- Верхній поріг оцінювання L_{up} , який обчислюється окремо для кожної речовини за алгоритмами з додатку 2 порядку здійснення державного моніторингу в галузі охорони атмосферного повітря [10];

- Нижній поріг оцінювання L_{low} , який обчислюється окремо для кожної речовини за алгоритмами з додатку 2 порядку здійснення державного моніторингу в галузі охорони атмосферного повітря [10].

Для аналізу використовуються тільки первинні дані за 5 останніх років, на основі яких визначаються ці порогові значення та здійснюється порівняння.

На рис. 2.7 наведено приклад такого порівняння результатів спостережень за концентрацією діоксиду сірки із верхнім та нижнім порогом оцінювання, в результаті якого виявлено постійне перевищення нижнього порогу оцінювання та одиничні випадки перевищення верхнього порогу оцінювання.

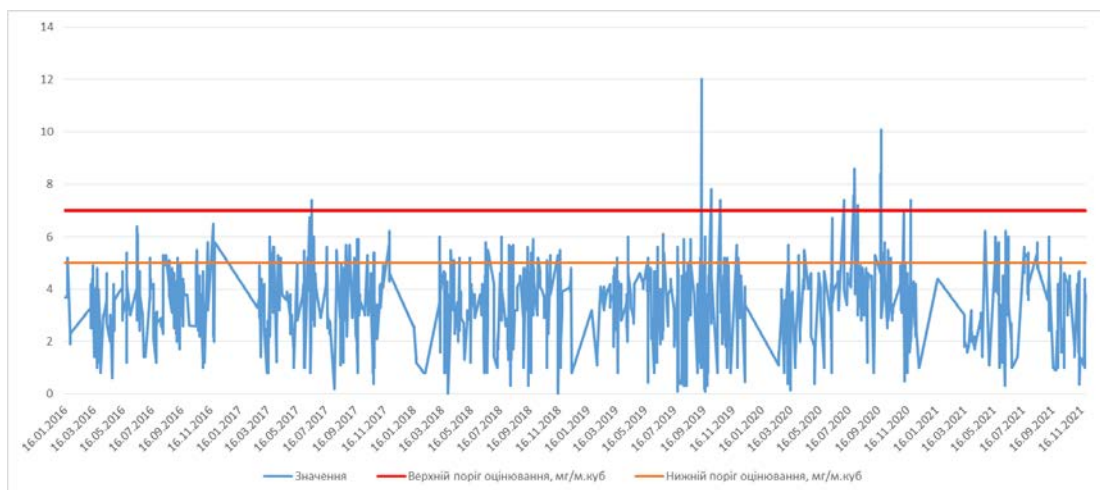


Рисунок 2.7 – Порівняння результатів спостережень діоксиду сірки з порогами оцінювання

2.3.2. Визначення мінімальної кількості постів і складу показників спостережень за вітчизняним та європейським законодавством

Згідно директиви 2008/50/ЄС [24], яка серед багатьох питань контролю за станом атмосферного повітря регламентує мінімальну кількість пунктів спостереження та відбору проб показників стану атмосферного повітря для їх подальшого порівняння з дозволеними рівнями. В Україні для її імплементації затверджено постанову Кабінету Міністрів України № 827 від 14 серпня 2019 року (поточна редакція від 11 травня 2024 року) [10] та наказ МВС України № 300 від 21 квітня 2021 року [12]. На їх основі розроблено алгоритм визначення мінімальної кількості постів і складу показників спостережень, який передбачає такі етапи:

1. Збір та аналіз існуючих даних.

Складання детальної мапи області з нанесенням усіх промислових підприємств, автомагістралей, житлових масивів та інших джерел забруднення.

Збір даних про існуючі пункти спостереження, їх технічне оснащення та якість отримуваних даних.

Аналіз існуючих нормативних документів, що регламентують моніторинг атмосферного повітря.

Оцінка відповідності існуючої системи моніторингу вимогам Наказу МВС України № 300 [12] та Постанови Кабінету Міністрів України № 827 [10].

2. Визначення критеріїв оптимізації.

Представленість джерел забруднення. Розміщення пунктів спостереження таким чином, щоб охопити всі основні джерела забруднення

(промислові підприємства, транспортні потоки тощо).

Густота населення. Пріоритетне розміщення пунктів у густонаселених районах та зонах рекреації.

Метеорологічні умови. Врахування рози вітрів, рельєфу місцевості та інших метеорологічних факторів, які впливають на розповсюдження забруднювальних речовин.

Технічні можливості. Оцінка наявності необхідного обладнання, кадрового забезпечення та фінансування.

Економічна ефективність. Вибір оптимального варіанту з точки зору витрат на обладнання, експлуатацію та обробку даних.

3. Моделювання розповсюдження забруднювальних речовин.

4. Застосування математичних моделей для прогнозування концентрації забруднювальних речовин в атмосферному повітрі при різних сценаріях розміщення пунктів спостереження.

Аналіз отриманих результатів та визначення зон з найбільш високими концентраціями забруднювальних речовин.

5. Розробка оптимальних схем розміщення пунктів спостереження:

Створення декількох варіантів схем розміщення пунктів з урахуванням визначених критеріїв.

Оцінка ефективності кожного варіанту за допомогою розроблених моделей.

Вибір оптимального варіанту, який забезпечує найбільш повну інформацію про стан атмосферного повітря при мінімальних витратах.

Розроблений алгоритм дозволяє оптимізувати систему моніторингу атмосферного повітря з урахуванням вимог чинного законодавства та забезпечує отримання достовірної інформації про стан атмосферного повітря в області.

Важливо зазначити, що оптимізація системи моніторингу є постійним процесом, який вимагає регулярного перегляду та коригування з урахуванням змін у промисловому виробництві, транспорті та інших факторах, що впливають на якість атмосферного повітря.

Для успішної реалізації алгоритму необхідна тісна співпраця органів державної влади, наукових установ та промислових підприємств.

Запропонований алгоритм надає загальну структуру для оптимізації системи моніторингу атмосферного повітря. Деталізація кожного етапу залежить від конкретних умов регіону та поставлених завдань. Графічне представлення цього алгоритму наведено на рисунку 2.8.

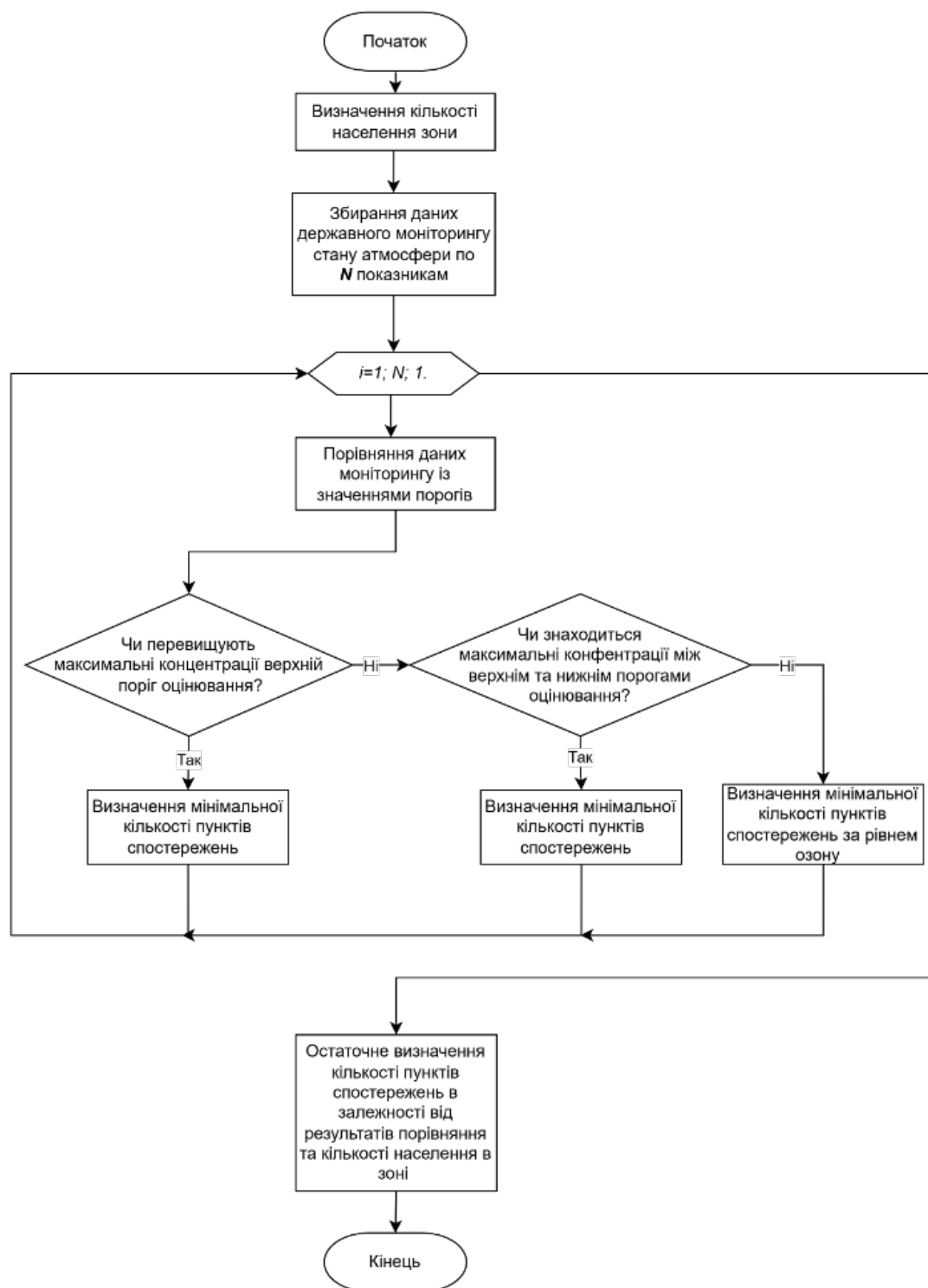


Рисунок 2.8 – Алгоритм оптимізації системи моніторингу на основі аналізу показників

Запропонований алгоритм був успішно випробуваний співавторами цієї монографії для оптимізації системи моніторингу на території зони «Вінницька» у 2021 році на замовлення Вінницької міськради. На

території зони «Вінницька» відсутні автоматизовані пости спостережень за якістю атмосферного повітря та опадів.

Відповідно до вимог наказу МВС України від 21.04.2021 р. № 300 «Порядок розміщення пунктів спостережень за забрудненням атмосферного повітря в зонах та агломераціях» [12] зони у відповідності з кількістю населення повинні мати відповідну мінімальну кількість пунктів спостережень з фіксованими вимірюваннями (рис. 2.9).

Населення агломерації чи зони (тисяч чоловік)	Якщо максимальні концентрації перевищують верхній поріг оцінювання				Якщо рівень максимальних концентрацій знаходиться між верхнім та нижнім порогами оцінювання				Мінімальна кількість пунктів спостережень для фіксованого вимірювання концентрацій озону		
	забруднювальна речовина, за винятком ТЧ	ТЧ (ТЧ ₁₀ + ТЧ _{2,5})	As, Cd, Ni	бенз(а) пірен	забруднювальна речовина, за винятком ТЧ	ТЧ (ТЧ ₁₀ + ТЧ _{2,5})	As, Cd, Ni	бенз(а) пірен	агломерації (міські та приміські)	інші зони (приміські та сільські)	сільські околиці
0-249	1	2	1	1	1	1	1	1		1	1 пункт на 50000 км ² як середня щільність по всіх зонах та 1 пункт на 25000 км ² для умов складного рельєфу
250-499	2	3			1	2			1	2	
500-749	2	3			1	2			2	2	
750-999	3	4	2	2	1	2	1	1	2	2	
1000-1499	4	6			2	3			3	3	
1500-1999	5	7			2	3			3	4	
2000-2749	6	8	2	3	3	4	1	1	4	5	
2750-3749	7	10			3	4			5	6	
3750-4749	8	11	3	4	3	6	2	2	один додатковий пункт на 2 млн мешканців	один додатковий пункт на 2 млн мешканців	
4750-5999	9	13	4	5	4	6	2	2			
>6000	10	15	5	5	4	7	2	2			

Рисунок 2.9 – Таблиця для визначення мінімальної кількості пунктів спостережень за забрудненням атмосферного повітря в зонах та агломераціях [12]

Як видно з рисунку 2.9, зона «Вінницька» з населенням 1 147,477 тис. жителів повинна мати мінімум 4 пункти спостережень з фіксованими вимірюваннями. Аналіз показників по концентрації пилу (і порівняння недиференційованого пилу з ГДК за даними державного моніторингу, і порівняння з верхнім порогом даних, визначеними в пункти

2.3.1, окремо по PM2.5 та PM10 за даними громадського моніторингу) показав наявність частих і значних перевищень, що означає необхідність спостережень за мінімально 6, а не 4 пунктами спостережень.

Головним недоліком цього алгоритму є відсутність пріоритетності щодо постів, яка могла б визначити послідовність закупівлі і монтажу цих постів, у разі браку фінансування на усі разом. Крім того, цей алгоритм встановлює мінімальну кількість постів спостережень, але їх може бути й більше, а тоді варто було б знати які і де слід встановлювати наступними. А тому на основі цього алгоритму був розроблений новий на основі методу аналізу ієрархій. Передусім, варто визначитись з системою критеріїв відбору постів, враховуючи вище наведені ключові аспекти, але врахувавши їх аномальні або максимальні показники.

2.3.3. Формування системи критеріїв для вибору параметрів місць постів спостережень за станом атмосферного повітря, з урахуванням екстремальних значень показників

Пропонується метод формування ієрархічної системи критеріїв для вибору параметрів місць постів спостережень за станом атмосферного повітря, з урахуванням екстремальних значень показників, які є характерними для регіону і наведені вище.

У додатку Г наведено результат аналізу невизначеності інформації про складові моделі DPSIR для стану атмосферного повітря Вінницької області, де враховано випадки, коли інформація є, але вона не зовсім точно відповідає вимогам [10-12] і усунути цю проблему неможливо.

Важливо враховувати 3 ключові групи факторів:

1) **D-P** – продуктивні сили, у тому числі різні антропогенні джерела (**D**), які впливають на стан атмосферного повітря через викидання різного роду забруднювальних речовин (**P**);

2) поточний стан атмосферного повітря (**S**), який визначається за даними мережі спостереження за станом атмосферного повітря;

3) вплив на населення та довкілля (**I**), що робить необхідним прийняття відповідних управлінських рішень (**R**).

Тому, важливо розташовувати станції моніторингу з урахуванням трьох основних видів критеріїв:

1. Знаходження неподалік від місць відомих аномально великих антропогенних джерел викидів – підприємства машинобудування, хімічного виробництва, металургійні підприємства, тощо;

2. Розташування у місцях, де раніше було виявлено перевищення показників якості атмосферного повітря відносно певних граничних значень або середніх максимальних;

3. Розташування неподалік від найбільш місць проживання особливо вразливих верств населення. До таких місць можуть відноситися санаторії, дитячі лікарні, дитячі табори, школи тощо.

Залучення експертів та аналіз рекомендацій нормативної бази Міндовкілля, з урахуванням наведених вище критеріїв, показав, що важливими є такі 5 критеріїв:

Φ_1 – близькість до найбільш впливових джерел забруднення атмосферного повітря (група Р-критеріїв – «Pressures»);

Φ_2 – частота та рівень перевищень усереднених та максимальних разових показників стану атмосферного повітря та атмосферних опадів (група S-критеріїв – «States»);

Φ_3 – близькість до стейкхолдерів (населення, у т.ч. вразливого, інвестиційних об'єктів тощо) (група I-критеріїв – «Impacts», інвестиції можна умовно розглядати як непрямі позитивні «наслідки» – у більш чистих регіонах більше інвестують, оскільки меншими будуть витрати на покращення здоров'я працівників, на фільтри у вентиляційних системах та ін.);

Φ_4 – рівень репрезентативності запроваджених керівних заходів (встановлені очисні системи, введений контроль за викидами пересувних транспортних засобів тощо) (група R-критеріїв – «Responses»);

Φ_5 – зручність інфраструктури, умов для обслуговування пунктів спостережень, цінність для рівномірного охоплення регіону, що спростить побудову карт забруднення в майбутньому та ін. (група допоміжних критеріїв).

Для кожного інтегрального критерія необхідно визначити декілька базових критеріїв (рис. 2.10).

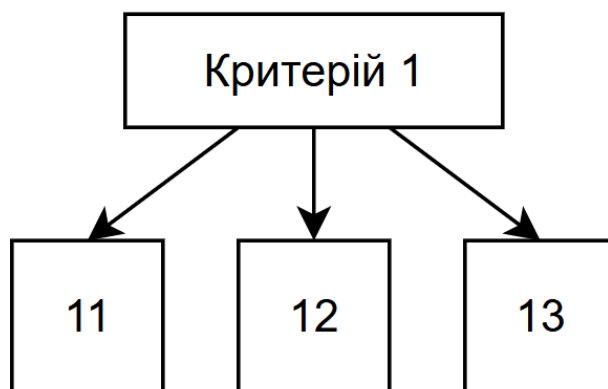


Рисунок 2.10 – Структура ієрархічного критерія вибору місця розміщення станції

У 1-му інтегральному критерії пріоритетного вибору місць розташування пунктів спостережень якості атмосферного повітря регіону базовими критеріями є такі:

Φ_{11} – сумарний обсяг стаціонарних джерел викидів підприємств в останній із наданих років зі списку найбільших забруднювачів повітря регіону у певному околі (наприклад у колі, площею 1/100 чи 1/1000 від площі регіону);

Φ_{12} – сумарний обсяг гранично допустимих викидів підприємств в останній із наданих років зі списку найбільших забруднювачів повітря регіону у певному околі;

Φ_{13} – кількість центрів виробничих майданчиків у певному околі;

Φ_{14} – близькість до автостради міжнародного чи регіонального значення або дороги з високою завантаженістю пересувними транспортними засобами у години «пік», передусім вантажними (експертно);

Φ_{15} – ступінь близькості від пункту до геометричного центра сусіднього населеного пункту або території з найбільшою кількістю стаціонарних джерел викидів.

У 2-му інтегральному критерії пріоритетного вибору місць розташування пунктів спостережень якості атмосферного повітря регіону базовими критеріями є такі:

Φ_{21} – відсоток перевищень середньодобової ГДК показника у точках у певному околі (наприклад у колі, площею 1/100 чи 1/1000 від площі регіону);

Φ_{22} – відсоток перевищень максимально разової ГДК показника у точках у певному околі;

Φ_{23} – відсоток перевищень нижнього порогу оцінювання показника у точках у певному околі;

Φ_{24} – відсоток перевищень верхнього порогу показника у певному околі;

Φ_{25} – відсоток перевищень ГДК за супутниковими даними та/або даними громадського моніторингу атмосферного повітря у певному околі

У 3-му інтегральному критерії пріоритетного вибору місць розташування пунктів спостережень якості атмосферного повітря регіону базовими критеріями є такі:

Φ_{31} – близькість до інтенсивної забудови багатоповерхових житлових масивів міст, км;

Φ_{32} – близькість до місць розташування найбільш вразливого населення: санаторіїв, лікарень, шкіл, дитсадків тощо (експертно);

Φ_{33} – близькість до наявних інвестиційних об'єктів, км;

Φ_{34} – близькість до потенційних інвестиційних об'єктів (експертно);

Φ_{35} – близькість до органів самоврядування (міськрад, сільрад) та інших місць, де часто зустрічають різні делегації, у т.ч. іноземні, км.

У 4-му інтегральному критерії пріоритетного вибору місць розташування пунктів спостережень якості атмосферного повітря регіону базовими критеріями є такі:

Φ_{41} – близькість до об'єктів, де нещодавно були встановлені очисні споруди та проведено інші заходи зі зниження забруднення атмосферного повітря, у т.ч. асфальтування доріг та інших дорожньо-транспортних робіт, заходів з оптимізації дорожнього руху тощо (експертно);

Φ_{42} – близькість до об'єктів, де планується встановити очисні споруди чи провести інші заходи зі зниження забруднення атмосферного повітря, у т.ч. асфальтування доріг та інших дорожньо-транспортних робіт, заходів з оптимізації дорожнього руху тощо (експертно);

Φ_{43} – близькість до місць, де запроваджено контроль за викидами пересувних джерел, за вагою вантажних засобів та ін., км;

Φ_{44} – близькість до об'єктів природно-заповідного фонду або сільської місцевості (для зон);

Φ_{45} – близькість до паркових зон чи зон з обмеженим рухом для вантажного транспорту.

У 5-му інтегральному критерії пріоритетного вибору місць розташування пунктів спостережень якості атмосферного повітря регіону базовими критеріями є такі:

Φ_{51} – зручність для обслуговування вимірювального обладнання пункту (експертно);

Φ_{52} – зручність інфраструктури для встановлення вимірювального обладнання пункту (експертно);

Φ_{53} – цінність для рівномірного охоплення регіону, що спростить побудову карт забруднення в майбутньому (експертно);

Φ_{54} – цінність для моніторингу за якістю атмосферного повітря регіону на думку органу, відповідального за моніторинг якості атмосферного повітря (експертно);

Φ_{55} – цінність для контролю за регулярними порушниками умов, які прописані в дозволах підприємств даного регіону, на думку органу, відповідального за екоінспекційний контроль якості атмосферного повітря (експертно).

2.3.4. Приклад застосування з використанням методу аналізу ієрархій

До визначених у попередньому розділі критеріїв та експертних оцінок застосовуємо стандартний алгоритм методу аналізу ієрархій [35, 36]. Є декілька алгоритмів його застосування. Пропонується використовувати алгоритм зі статті [37]. Як відомо, за цим методом знаходять максимальне дійсне додатне власне число $\lambda_{\max}$ кожної C -матриці та власні вектори w^j для нього ($j = 0 \dots n$) та розв'язують систему матричних рівнянь виду [37]

$$\left[C^{(j)} - \lambda_{\max}^{(j)} I \right] w^{(j)} = 0, \text{ де } j = 1 \dots n. \quad (2.7)$$

При цьому, необхідно замінити одне з рівнянь на таке [37]:

$$w_1^j + w_2^j + \dots + w_n^j = 1. \quad (2.8)$$

Розв'язавши систему рівнянь, ми отримуємо ваги критеріїв $w_1^{(j)}$, де $j = 1 \dots n$ та можемо використати їх для обчислення значень інтегральних факторів та кожного базового фактору за допомогою наступних формул [37]:

$$\begin{aligned} K_j &= w_j^{(0)} K_{nn}^*, \text{ де } j = 1 \dots n, \\ K_{ji} &= w_j^{(0)} w_i^{(j)} K_{nn}^*, \text{ де } j = 1 \dots n, i = 1 \dots n. \end{aligned} \quad (2.9)$$

В цілому, блок-схема алгоритму запропонованої технології наведена на рисунку 2.11.

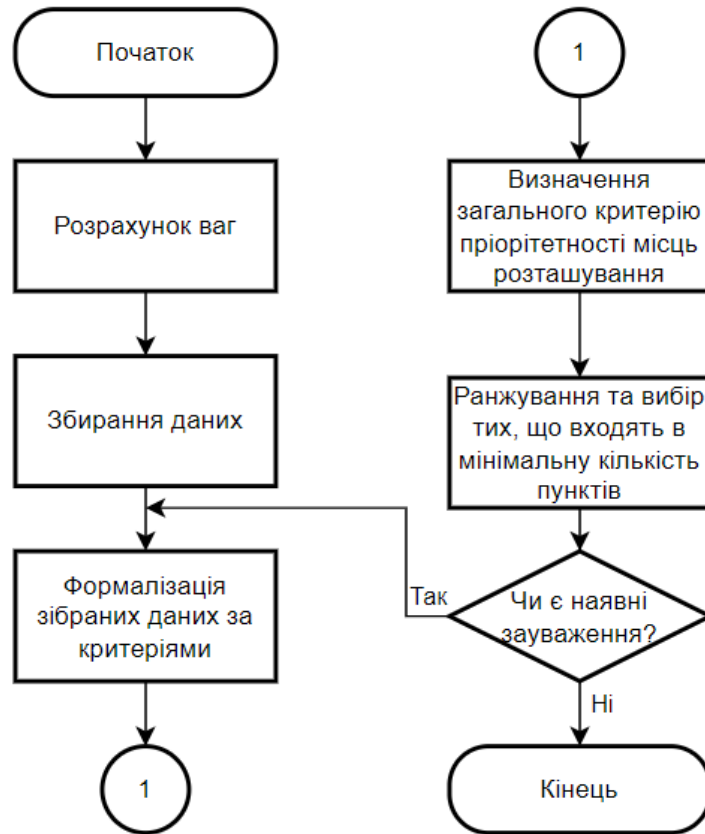


Рисунок 2.11 – Блок-схема алгоритму запропонованої технології проектування мережі спостережень якості атмосферного повітря регіону на основі методу аналізу ієрархій [38]

Наведемо приклад обрахунків згідно описаного методу. Експерти проводять попарне порівняння критерії та визначають за допомогою оцінок наскільки один критерій переважає над іншим. З цих оцінок формується матриця, як це показано в таблиці 2.1.

Таблиця 2.1 – Приклад експертних оцінок по визначених інтегральних критеріях [38, 39]

Критерії	Φ_1	Φ_2	Φ_3	Φ_4	Φ_5
Φ_1	1	2	5	6	6
Φ_2	$\frac{1}{2}$	1	3	4	4
Φ_3	$\frac{1}{5}$	$\frac{1}{3}$	1	4	2
Φ_4	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{4}$	1	2
Φ_5	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{2}$	1

Використовуючи формули 2.7 і 2.8 формується система рівнянь (рис 2.12).

$$\begin{aligned} -4.258*w_1 + 2.0*w_2 + 5.0*w_3 + 6.0*w_4 + 6.0*w_5 &= 0 \\ 0.5*w_1 + -4.258*w_2 + 3.0*w_3 + 4.0*w_4 + 4.0*w_5 &= 0 \\ 0.2*w_1 + 0.333*w_2 + -4.258*w_3 + 4.0*w_4 + 2.0*w_5 &= 0 \\ 0.167*w_1 + 0.25*w_2 + 0.25*w_3 + -4.258*w_4 + 2.0*w_5 &= 0 \\ 0.167*w_1 + 0.25*w_2 + 0.5*w_3 + 0.5*w_4 + -4.258*w_5 &= 0 \end{aligned}$$

Рисунок 2.12 – Приклад системи рівнянь у разі застосування методу аналізу ієрархій [38, 39]

Розв'язавши цю систему за допомогою формули 2.10, ми отримуємо ваги кожного критерію (рис 2.13).

$$\begin{aligned} K_1 &= w_{0_1} * K_{nn} = 0.466 * K_{nn} \\ K_2 &= w_{0_2} * K_{nn} = 0.27 * K_{nn} \\ K_3 &= w_{0_3} * K_{nn} = 0.136 * K_{nn} \\ K_4 &= w_{0_4} * K_{nn} = 0.069 * K_{nn} \\ K_5 &= w_{0_5} * K_{nn} = 0.058 * K_{nn} \end{aligned}$$

Рисунок 2.13 – Приклади обрахованих ваг інтегральних критеріїв [38, 39]

Метод застосовується для кожного показника якості атмосферного повітря та атмосферних опадів окремо. Спочатку збираємо дані, потім за кожним критерієм визначаємо 3-5 оптимальних місць (наприклад, за Φ_{11} – де найбільше стаціонарних джерел викидів, за Φ_{32} – де найбільше лікарень, санаторіїв, шкіл і дитсадків, за Φ_{51} та Φ_{52} – у пунктах, де вже є стаціонарні пости, наприклад ДСНС України, або місця, де регулярно проводить вимірювання МОЗ, та ін.), формуємо єдиний список з усіх цих місць, проводимо розрахунки, ранжуємо усі місця по даному показнику, відбираємо з них першочергові для встановлення. Потім можна додати ранги цих місць по усіх показниках і вибрати найбільш

пріоритетні місця одночасно за усіма показниками, але можна і не додавати. Іноді краще в одному місці вимірювати пил, в другому – важкі метали, у третьому – озон (давач на озон – порівняно дешевий, його можна й в усіх пріоритетних місцях вимірювати), у четвертому – якість атмосферних опадів. Це вже – на розсуд проектувальників та комісії зони чи агломерації.

Таким чином, застосування запропонованого підходу формування критеріїв оптимізації мережі спостереження дозволяє розкрити невизначеність U_{T1} . Під час формування цих критеріїв використовуються порогові значення показників, згідно офіційних директив, що дозволяє розкрити невизначеність U_{T3} та U_{T5} .

2.4. Висновки до розділу 2

Виходячи з аналізу проведеного у розділі 1, було сформовано математичну постановку задачі управління станом атмосферного повітря з використанням методології системного аналізу.

Описано метод пошуку можливих джерел підвищеного забруднення на основі аналізу аномалій, який дозволяє за наявними даними моніторингу стану атмосферного повітря локалізувати у часі та просторі джерела, які є причиною зростаючого рівня забруднення атмосферного повітря. Це, у свою чергу, дозволило розкрити один з видів невизначеності.

Запропоновано метод формування ієрархічної системи критеріїв для оптимізації мережі моніторингу атмосферного повітря, яка містить етап розроблення інтегральних критеріїв, що, в першу чергу, базуються на аномальних та пікових значеннях стану відповідного природного середовища. Далі, застосовуючи метод аналізу ієрархій, проводиться вибір оптимальних місць для встановлення станцій моніторингу, що дозволяє ефективно використовувати ресурси та отримувати при цьому корисну інформацію, необхідну для прийняття обґрунтованих керівних рішень.

РОЗДІЛ 3

ПРОГНОЗУВАННЯ СТАНУ АТМОСФЕРНОГО ПОВІТРЯ З УРАХУВАННЯМ ПІКОВИХ ЗНАЧЕНЬ

3.1. Метод ідентифікації параметрів сезонності часового ряду

В підрозділі 1.6 було визначено, що існують види невизначеності, пов'язані з побудовою ефективної інтелектуальної моделі прогнозування стану атмосферного повітря, а саме:

- невизначеність щодо значень параметрів моделей (U_{P1});
- невизначеність щодо параметрів моделей, які описують вплив сезонної складової: період, вікно впливу, коефіцієнт регуляризації або сила впливу (U_{P2});
- вибір того, яка саме сезонна складова або їх сукупність мають місце (U_{T6}).

Для розкриття цих видів невизначеності, необхідно розглянути класичні підходи розв'язання таких задач. Також варто запропонувати метод або набір методів, які дозволять проаналізувати наявні дані моніторингу стану атмосферного повітря та визначити основні параметри інтелектуальної моделі прогнозування. Важливими параметрами періодичних часових рядів є параметри їх тренду, сезонних складових та аномалій. Задача визначення тренду часового ряду має багато універсальних методів розв'язання, проте, ідентифікація одночасно параметрів різних видів сезонності та аномалій різної природи у різні часові проміжки є складною задачею, яка не має універсального вирішення.

Розглянемо декілька класичних підходів до вирішення задачі.

3.1.1. Дослідження впливу параметрів моделі Prophet на точність прогнозування стану атмосферного повітря

Модель Prophet, яку було детально описано в підрозділі 1.4, дозволяє налаштовувати багато параметрів і, тим самим, моделювати часовий ряд, враховуючи експертні знання та певні особливості самого ряду. Одним з таких параметрів є набір «holidays», що дозволяє вказати перелік відомих дат аномальних впливів, наприклад, природних або антропогенних катастроф. Використання цього параметру дозволяє зменшити вплив відповідних аномальних дат на точність роботи моделі при прогнозуванні значень показників стану атмосферного повітря.

Іншим важливим параметром моделі Prophet є параметр регуляризації «prior_scale», що дозволяє контролювати наскільки сильно вказані

аномальні дати впливають на прогноз [40]. Застосування великих значення вказує моделі більше реагувати на аномальні дати, що, відповідно, збільшує їх вплив на прогноз.

Варто перевірити як згадані налаштування впливають на якість роботи моделі. Для цього, створимо програму на Python для пошуку аномальних значень за рахунок порівняння їх з пороговим значенням, яке визначається певним квантилем. На рисунку 3.1 наведено часовий ряд показника пилу PM10, отриманого від мережі громадського моніторингу EcoCity, яка була згадана раніше в пункті 1.1.4. На рисунку червоними точками відмічені аномальні дати, які були визначені за рахунок порівняння усіх точок ряду з квантилем 95%.

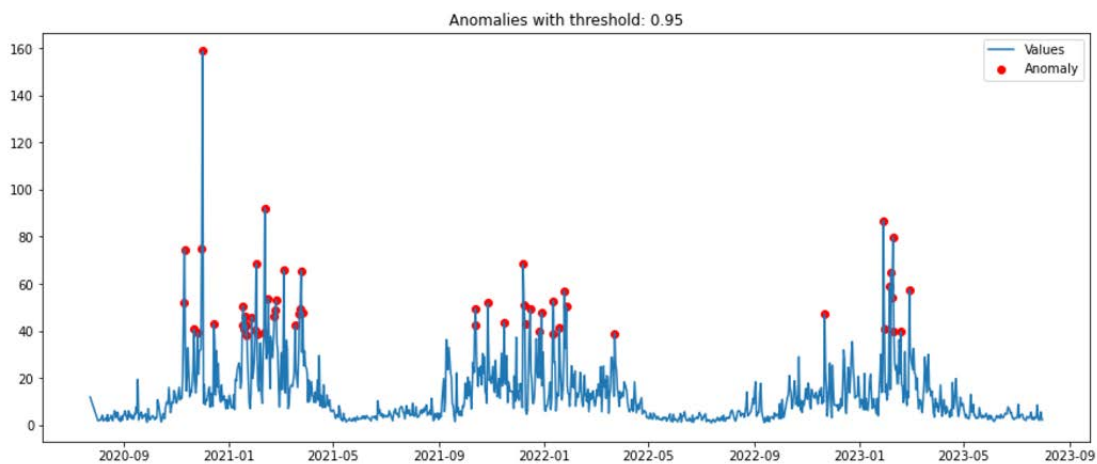


Рисунок 3.1 – Приклад показника пилу PM10, отриманого від мережі громадського моніторингу EcoCity [40]

Обираємо декілька значень для параметру «prior_scale» та декілька значень квантилів. Для часового ряду обраного параметру видаляються аномальні дані, використовуючи відповідний квантиль, і будується модель Prophet. Кожна модель будується з налаштуванням денної сезонності за замовчуванням та параметром «holidays», який є набором аномальних даних, обраних раніше та відповідним параметром «prior_scale». Ці кроки повторюються для усіх можливих комбінацій використання цих двох параметрів. Для кожного набору тестових даних робиться прогноз та обраховується показник точності. Приклад результату роботи моделі Prophet для часового ряду з аномаліями, визначеними по квантилю 95% та з значенням «prior_scale» 10, наведено на рисунку 3.2.

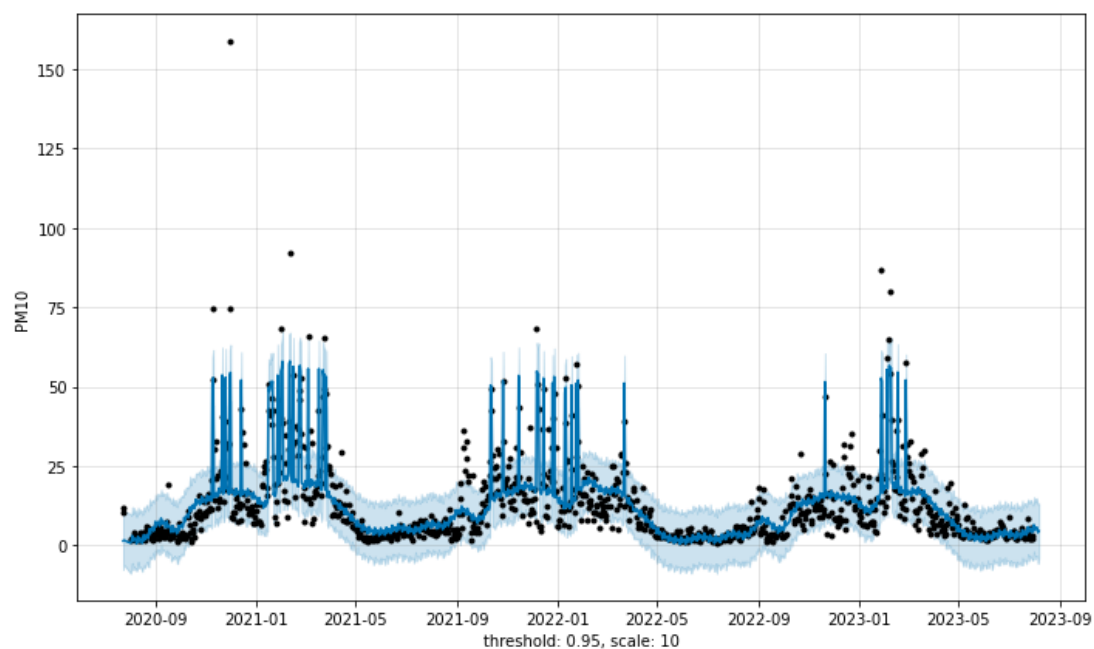


Рисунок 3.2 –Результат роботи моделі Prophet для даних, наведених на рис 3.1 [40]

Результат роботи моделі Prophet порівнюється з реальними значеннями ряду, використовуючи коефіцієнт детермінації (зазвичай, позначається як «r2_score») Результати для усіх комбінацій вхідних параметрів наведені на рисунку 3.3.

	threshold	scale	r2
0	0.50	1.0	51.72
1	0.50	5.0	51.84
2	0.50	10.0	51.91
3	0.50	15.0	51.88
4	0.50	20.0	51.89
5	0.50	30.0	51.85
6	0.70	1.0	64.38
7	0.70	5.0	64.48
8	0.70	10.0	64.51
9	0.70	15.0	64.49
10	0.70	20.0	64.49
11	0.70	30.0	64.50
12	0.85	1.0	70.37

	threshold	scale	r2
12	0.85	1.0	70.37
13	0.85	5.0	70.43
14	0.85	10.0	70.44
15	0.85	15.0	70.43
16	0.85	20.0	70.43
17	0.85	30.0	70.46
18	0.95	1.0	66.20
19	0.95	5.0	66.27
20	0.95	10.0	66.27
21	0.95	15.0	66.27
22	0.95	20.0	66.26
23	0.95	30.0	66.26

Рисунок 3.3 – Результати перевірки впливу вхідних параметрів на точність передбачення [40]

Аналізуючи отримані результати, можна зробити висновок, що зміна параметру `prior_scale` не мала істотного впливу на прогнозування певного квантиля, а сама точність передбачення збільшується в межах використання квантиля P70-P85, але при подальшому збільшенні точність починає погіршуватися.

3.1.2. Ідентифікація порядку ряду Фур'є для показника стану атмосферного повітря

У моделі Prophet для моделювання сезонних компонентів використовується ряд Фур'є, згаданий раніше в підрозділі 1.4. Для пошуку порядку ряду Фур'є, зазвичай, використовують наступні підходи [37]: частотний аналіз, аналіз спектра та перебір усіх можливих значень, наприклад від 3 до 12 чи 20, причому останній підхід є більш популярним. Однак, як правило, визначення порядку ряду Фур'є здійснюється не самостійно. Разом із ним ще визначаються й інші параметри моделі, а тоді кількість комбінацій для повного перебору зростає дуже сильно – часто їх взагалі не можна перебрати усі, коли інші параметри є дробовими числами, а тому потрібні більш точні методи, які, на відміну від інших, добре працюють з сильно зашумленими даними.

Відповідно, корисним є розроблення підходу до визначення порядку Фур'є у моделі Prophet, що дозволить моделювати сезонну складову часового ряду за умов його значної зашумленості.

Для вирішення поставленої задачі пропонується наступний підхід [41]:

1) Синтезуємо ряд Фур'є для заданого порядку n (решта змінних задається фіксованими):

$$f(x) = \sum_{i=1}^n a_i \cos ix + b_i \sin ix. \quad (3.1)$$

2) На отриманий набір значень накладаємо шум із певною дисперсією d_1 ;

3) Використовуючи синтезовані на кроці 2 дані, будуємо модель Prophet. При цьому, ми відключаємо налаштування сезонності за замовчуванням та заміняємо на заданий порядок n ряду Фур'є;

4) Робимо прогноз для створеної моделі, а на отримані значення додатково накладаємо шум із заданою дисперсією d_2 ;

5) Повторюємо кроки 1-4 для різних наборів вхідних параметрів (порядок ряду і дисперсія шумів) для генерування набору тренувальних даних;

6) Використовуємо тренувальні дані, отримані на кроці 5, для ідентифікації порядку ряду з використанням методів машинного навчання.

Було реалізовано функцію, яка, перебираючи в циклі різні значення порядку ряду Фур'є та дисперсії шуму, генерує дані моделлю Prophet. Приклад роботи даної функції зображено на рисунку 3.4.

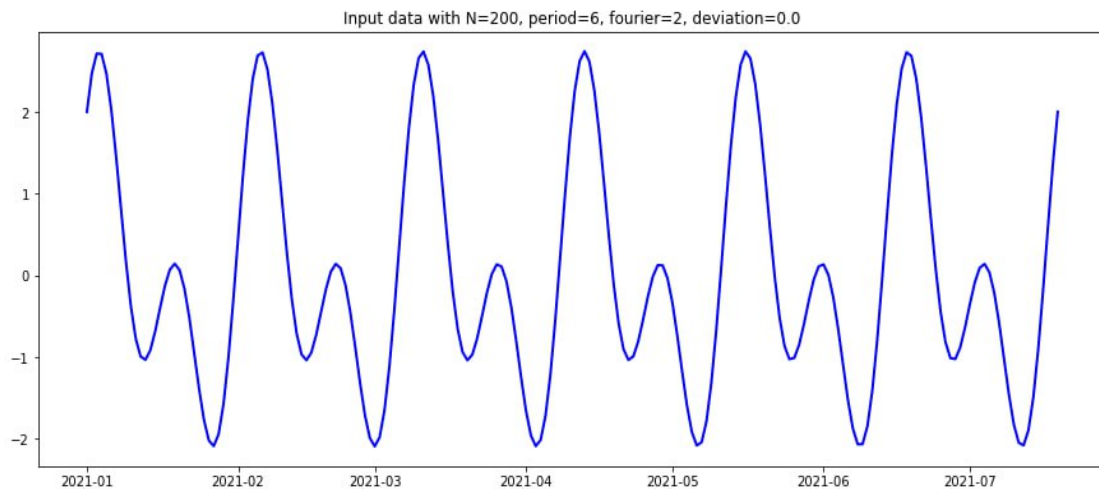


Рисунок 3.4 – Приклад роботи функції для генерування даних за допомогою моделі Prophet [41]

Використовуючи описану функцію, було отримано набір тренувальних даних із більше як 800 різних комбінацій порядку ряду Фур'є та дисперсії шуму (рис. 3.5).

value_1	value_2	value_3	value_4	value_5	value_6	...	value_1
1.212929	1.033774	1.490677	1.588684	1.327723	1.247872	...	-0.438
0.839207	1.214710	1.414985	1.412930	1.583367	1.651833	...	-0.542
0.881902	1.257847	1.499206	0.962251	1.067706	0.983035	...	-0.245
1.235705	0.767356	0.820297	1.253692	1.766596	1.389924	...	-0.892
0.903866	1.106160	1.049172	1.790335	1.210122	1.248544	...	-0.434
...	...	...	...	...	...	...	...
12.785682	5.139152	0.473676	2.031674	-0.133421	-0.055842	...	-1.828
13.039057	4.370853	-0.426115	1.719395	-0.604905	-0.593807	...	-0.983
12.363012	4.422465	0.239665	2.585668	0.533538	-0.348935	...	0.1628
13.145853	4.489238	-0.905436	0.917066	1.230329	-0.708511	...	-0.987
12.532139	4.957577	-1.328651	1.285472	-0.185101	-0.434669	...	-0.967

Рисунок 3.5 – Сформовані тренувальні дані [41]

Тренувальні дані були використані для ідентифікації порядку ряду Фур'є з використанням методу RandomForestRegressor. Для тестування було використано вибірково взяті 20% даних, які не використовувались для тренування моделі. Відносна похибка визначення порядку Фур'є склала 2-3%, що є дуже гарним показником. Зауважимо, що у разі збільшення дисперсії шуму, ця похибка суттєво зростає, тому цей метод має обмежене застосування.

Загалом, аналіз показав, що традиційні підходи не є достатньо ефективними та варто розробити новий метод для вирішення поставленої задачі.

3.1.3. Розроблення нового методу ідентифікації параметрів сезонності

Для розв'язання поставленої задачі пропонується новий метод в якому максимально автоматизується кожен етап традиційного розв'язання з використанням автоматичних процедур та ряду критеріїв.

Етап 1. Визначення кількості та періоду можливих видів сезонності.

У програмі [42] одним зі співавторів було запропоновано будувати графік усіх варіантів відношень амплітуд $S_y(P)$ з періодом P від 1 до 50% від кількості усіх значень ряду:

$$S_y(P) = \frac{S(P)}{Y}, \quad (3.2)$$

де $P = 1 \dots 0.5N$, N – кількість значень ряду.

Приклад такого графіку (назвемо його «декомпозиційною кривою») $S_y(P)$ наведено на рис. 3.6.

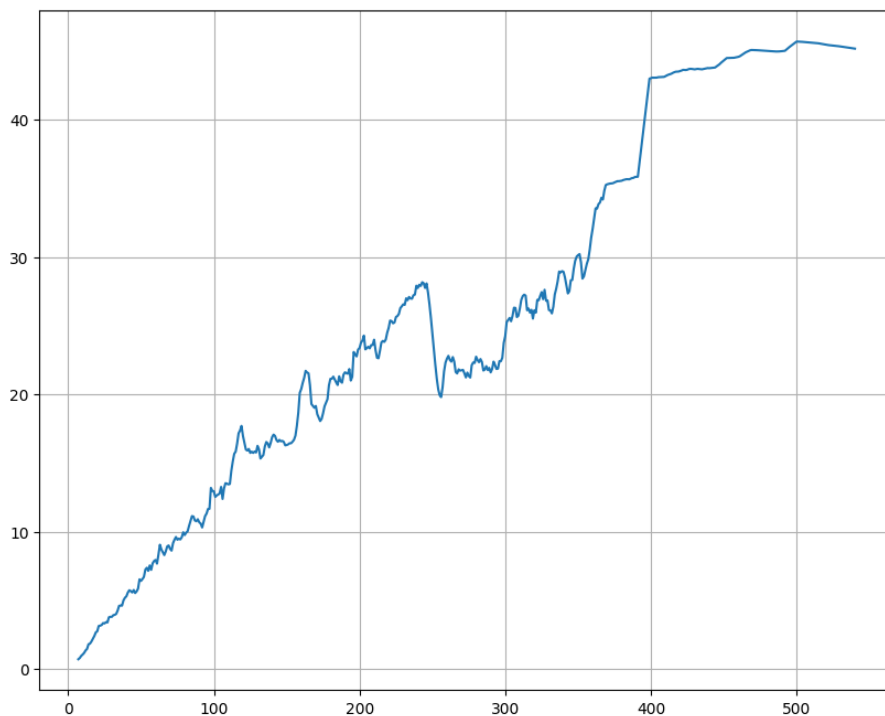


Рисунок 3.6 – Декомпозиційна крива часового ряду значення РМ10 за даними станції № 650 мережі громадського моніторингу атмосферного повітря EсоCity

за 2020-2023 рр. [25, 43]

Пропонується згладити декомпозиційну криву та знайти усі номери кроків, яким відповідають локальні максимуми, тобто місця, де зростання значення $S_y(P)$ змінюється на зменшення, це і будуть періоди P_i :

$$\begin{aligned} P_i : S'_Y(P_i) &= 0, \\ S''_Y(P_i) &< 0, \text{ де } i=1...\phi. \end{aligned} \quad (3.3)$$

Етап 2. Визначення параметрів видів сезонності.

Значний досвід авторів протягом 2020-2024 років роботи з моделлю Prophet показав, що у випадку, коли у моделі поєднується декілька видів сезонності з різним періодом, то для мінімізації ризику перенавчання ефективними є такі прийоми [25]:

- сезонність з найбільшим періодом варто описувати рядом Фур'є найменшого порядку $n_{\min}$ і з найбільшою регуляризацією $R_{s\min}$ (найменше значення саме забезпечує найбільші обмеження на коефіцієнти, тобто – найбільшу регуляризацію), тобто вона має бути найбільш гладкою;

- сезонність з найменшим періодом варто описувати рядом Фур'є найбільшого порядку $n_{\max}$ і з найменшою регуляризацією $R_{s\max}$, тобто вона має бути найбільш негладкою, тобто найкраще припасовуватись до ряду;

- порядок n_i ряду Фур'є варто брати між $n_{\min} = 2$ і $n_{\max} = 12$, інші межі (наприклад, 0, 1 чи 20) менш ефективні;

- регуляризацію сезонності варто брати між $R_{s\min} = 0.1$ і $R_{s\max} = 0.5$, інші межі (наприклад, 0,01, 0,05 чи 0,9) менш ефективні;

- у разі вибору між різними варіантами сезонних складових першими слід враховувати складові з більшим періодом, послідовно додаючи ті, в яких він – менший.

Ці прийоми дають можливість вивести універсальні співвідношення для розрахунку параметрів усіх видів сезонності. Зокрема, пропонуються такі нові співвідношення для i -ої складової, враховуючи, що період P_0 є найбільшим значенням [25]:

$$n_i = \frac{n_{\min} - n_{\max}}{\phi - 1} i + n_{\max}, \quad (3.4)$$

$$R_{si} = \frac{R_{s\max} - R_{s\min}}{\phi - 1} i + R_{s\min}, \text{ де } i=1...(\phi - 1). \quad (3.5)$$

Етап 3. Визначення параметрів аномалій.

Для забезпечення більшої універсальності методу пропонується

спрощення, шляхом відмови від різних видів аномалій. Крім того, пропонується під видами аномалій розрізняти не різні за формалізацією чи природою аномалії, а – аномалії у різні періоди ряду [25]:

$$y(t) = g(R_g, t) + \prod_{i=1}^{\phi} s(R_{si}, P_i, n_i, t) + h\left(\bigcup_{i=1}^{\phi} \bigcup_{j=1}^{\phi} H_{ij}(R_h, t)\right) + \varepsilon_t, \quad (3.6)$$

тобто для кожного періоду, визначеного на попередньому етапі, пропонується робити послідовні вибірки по одному періоду, скільки їх має місце у ряді (1-й, 2-й, 3-й..., останній може бути неповним) і в кожному окремо шукати аномалії, а потім усі об'єднувати в єдину множину. Усі аномалії брати з нульовим вікном, тобто вважати, що вони не впливали до чи після, а якщо впливали, то це – інші аномалії, які треба враховувати окремо. Регуляризацію для усіх аномалій пропонується брати однакову R_h , наприклад, 1 або 10, або враховувати як гіперпараметр.

Однак, крім цього, потрібний критерій для перевірки чи дійсно вибрані у такий спосіб точки є аномаліями. Тому пропонується для кожної вибірки для другої складової виразу (3.6) перевіряти нульову гіпотезу:

H_0 – нульова гіпотеза про те, що відібрана множина значень не є аномаліями відносно інших значень вибірки;

H_1 – альтернативна гіпотеза про те, що відібрана множина значень є аномаліями відносно інших значень вибірки.

У разі, якщо нульова гіпотеза підтверджується, тоді варто результат, тобто не враховувати знайдені точки як аномальні.

Етап 4. Ідентифікація гіперпараметрів.

На етапах 1-3 визначаються більшість параметрів, але частина має залишатись гіперпараметрами для формування множини можливих моделей і вибору оптимальної за точністю. Критерієм точності може бути відносна похибка або відомі метрики: середня абсолютна похибка (англ. «Mean absolute error» або скорочено – MAE), середньоквадратична похибка (англ. «Mean squared error» або скорочено – MSE) та коефіцієнт детермінації (зазвичай, позначається як «r2_score», або «R2» чи «r2») [26].

Гіперпараметрами моделі (3.6) можуть бути такі:

- 1) регуляризація тренду R_g (має бути побільше – на рівні $R_{s\min}$);
- 2) кількість ϕ видів сезонності, які варто брати до уваги з числа відібраних на етапі 1;
- 3) граничні значення порядку Фур'є $n_{\min}$ і $n_{\max}$ та регуляризації $R_{s\min}$ і $R_{s\max}$;
- 4) регуляризація аномальної складової R_h .

Якщо ж усі їх задати числами, тоді варто вибирати між моделями, які враховують різну кількість комбінацій видів сезонності.

Блок-схема алгоритму запропонованого методу наведена на рисунку 3.7.

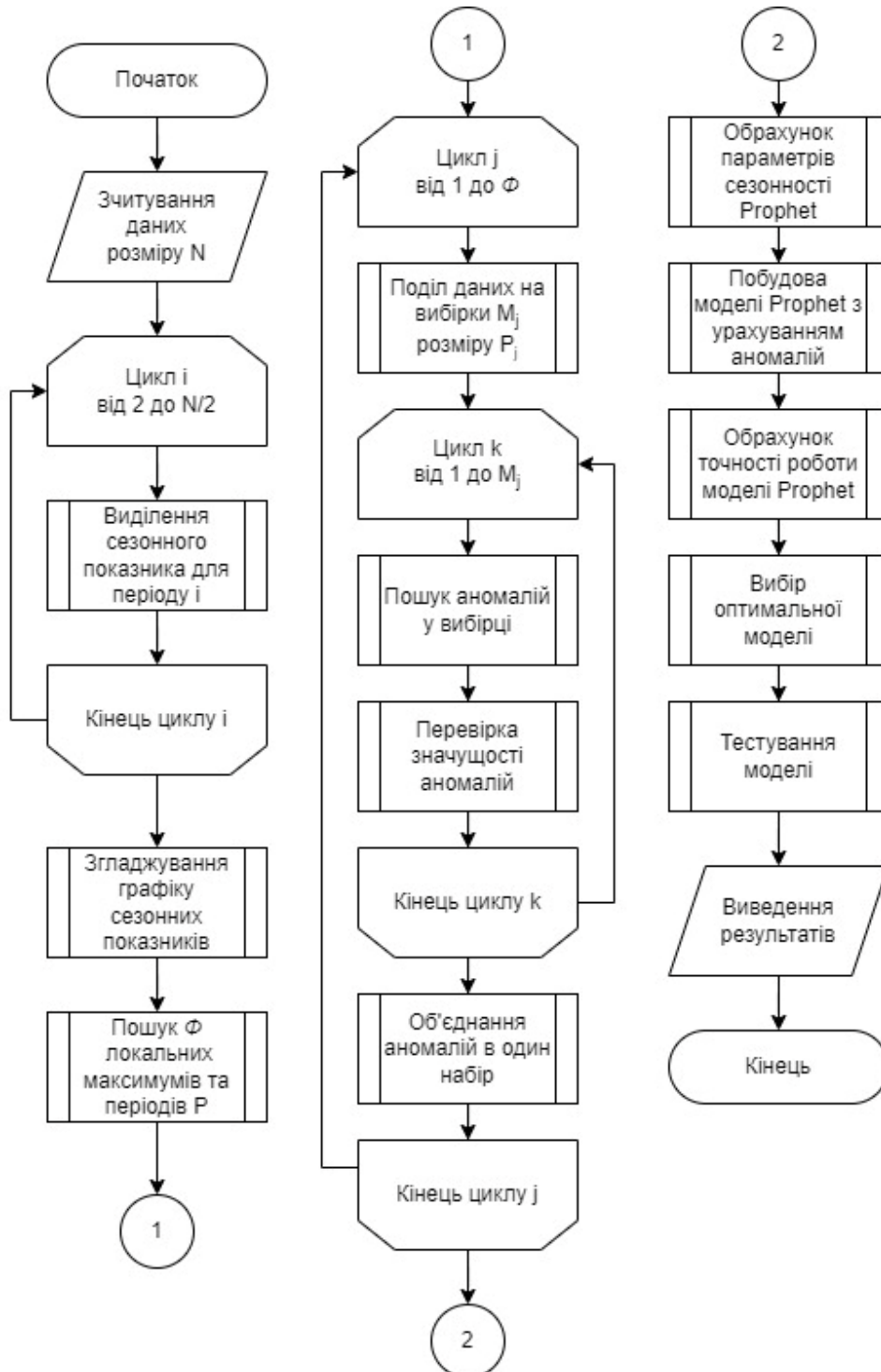


Рисунок 3.7 – Блок-схема алгоритму методу ідентифікації параметрів гармонік та аномалій періодичного часового ряду [25]

Для випробування даного методу використаємо дані показник пилу PM10 (мікроскопічні тверді частинки) за 3 роки (рис. 3.8), отриманого від мережі громадського моніторингу EcoCity, описаної раніше у пункті 1.1.4.

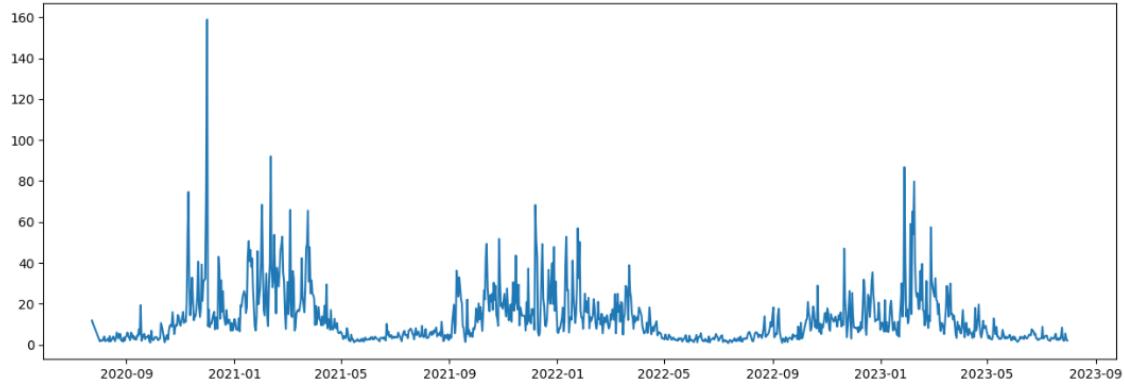


Рисунок 3.8 – Дані показника PM10 станції у смт Турбів Вінницького району [25, 43]

Застосуємо запропонований метод.

Етап 1. Декомпозиційний графік для ряду на рис. 3.8 було наведено на рис. 3.6. На рис. 3.9 наведено результат його згладжування фільтром Баттерворта з використанням Python-бібліотеки `scipy.signal`. Пошук локальних максимумів методом `find_peaks` бібліотеки `scipy.signal` дозволив з'ясувати, що $\phi = 3$: $P_1 = 230$, $P_2 = 419$, $P_3 = 507$ діб.

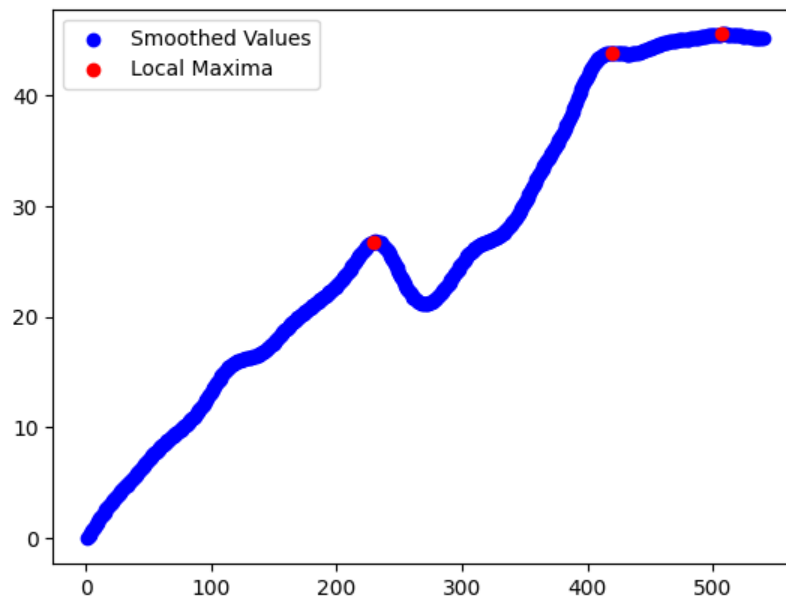


Рисунок 3.9 – Локальні максимуми для графіка з рис. 3.8, згладженого методом Баттерворта [25, 43]

Етап 2. Визначимо параметри сезонності з 3-ма різними періодами. Використовуючи формули (3.4) та (3.5), обчислимо інші параметри цих

сезонних складових (рис. 3.10).

```
period: 507, prior_scale: 0.1, fourier_order: 2
period: 419, prior_scale: 0.3, fourier_order: 7
period: 230, prior_scale: 0.5, fourier_order: 12
```

Рисунок 3.10 – Налаштування сезонності для моделі Prophet [25, 43]

Етап 3. Визначаємо параметри аномалій. Вхідний набір даних розділяється на рівні частини, де розмір частини відповідає визначеному на етапі 1 періоду. І ця операція повторюється для кожного виду сезонності окремо. Розглянемо для прикладу сезонну складову з періодом P_2 . Оскільки $P_2 = 419$, а загальна кількість даних складає 1080, то ми отримуємо дві вибірки розміром 419 та одну неповну вибірку з 186 даними. Для кожної з вибірок, застосовуючи метод Isolation Forest на основі дерев рішень, визначаємо аномалії, а тоді – перевіряємо нульову гіпотезу щодо того, що ці значення не є аномальними. В нашому випадку для перших двох вибірок нульова гіпотеза не підтвердилась і, відповідно, вважаємо аномалії на цих двох вибірках статистично значущими. Відповідно, як було зазначено вище, ці аномалії додаються в загальну вибірку усіх аномалій для відповідної періодичної складової. Результат пошуку аномалій зображено на рисунку 3.11, де червоними точками позначені відповідні елементи графіку показника PM10, які було ідентифіковано як аномалії. Для частини знайдених потенційних аномалій нульова гіпотеза підтвердилась, а тому вони не були додані до множини аномальних.

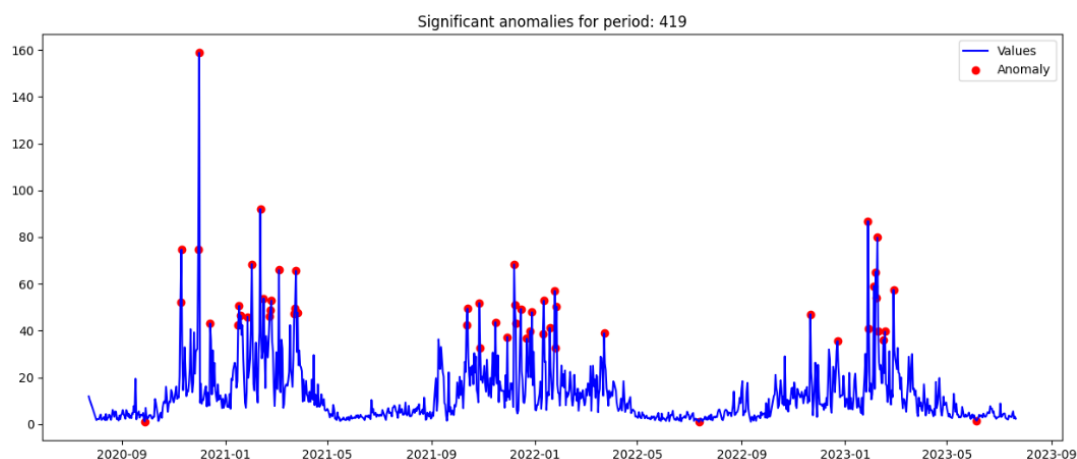


Рисунок 3.11 – Результат пошуку аномалій показника PM10 для періоду 419 [25, 43]

Етап 4. Сформувавши усі параметри на попередніх етапах, будуємо модель Prophet та використовуємо метрики, зазначені раніше, обраховуємо точність роботи моделі Prophet для різних варіантів (рис. 3.12):

Варіант 0 «Without default seasonality». Модель FB Prophet з параметрами за замовчуванням, але без сезонних складових.

Варіант 1. Варіант 0 з 3-ма видами сезонності (англ. «default seasonality»), які ідентифікуються за замовчуванням (добова, тижнева та річна).

Варіант 2. Варіант 0 з 3-ма визначеними на етапі 1 сезонними складовими одночасно ($\phi = 3$ з періодами $P_1 = 230, P_2 = 419, P_3 = 507$ діб) (англ. «custom seasonality»), але без аномалій («no holidays», оскільки в моделі FB Prophet аномалії прийнято умовно називати «свята»).

Варіант 3. Варіант 2 з усіма аномаліями («all anomalies»), визначеними у вибірках, розміром 230 діб.

Варіант 4. Варіант 3, але тільки зі статистично значущими аномаліями (англ. «significant anomalies»).

Варіант 5. Варіант 2 з усіма аномаліями, визначеними у вибірках, розміром 419 діб.

Варіант 6. Варіант 5, але тільки зі статистично значущими аномаліями.

Варіант 7. Варіант 2 з усіма аномаліями, визначеними у вибірках, розміром 507 діб.

Варіант 8. Варіант 7, але тільки зі статистично значущими аномаліями.

	name	r2	mean_squared_error	mean_absolute_error
0	Without default seasonality	0.264846	126.149253	7.053936
1	Default seasonality	0.401954	102.622188	6.125823
2	Custom seasonality + no holidays	0.431548	97.543824	5.901587
3	Period: 230 + all anomalies + custom seasonality	0.673193	56.078709	4.751458
4	Period: 230 + significant anomalies + custom s...	0.673193	56.078709	4.751458
5	Period: 419 + all anomalies + custom seasonality	0.699386	51.584003	4.647670
6	Period: 419 + significant anomalies + custom s...	0.699386	51.584003	4.647670
7	Period: 507 + all anomalies + custom seasonality	0.651996	59.715948	4.904736
8	Period: 507 + significant anomalies + custom s...	0.669741	56.671101	4.824506

Рисунок 3.12 – Таблиця порівняння точності апроксимації моделі Prophet для різних варіантів її архітектури (кількості сезонних складових) і параметрів [25, 43]

З рис. 3.12 видно, що для двох із трьох видів сезонності точність між статистично значущими аномаліями та усіма наявними аномаліями однакова. Причиною цього стало те, що перевірка нульової гіпотези визначила на усіх вибірках цих періодів аномалії суттєвими, відповідно для них ці два набори аномалій – абсолютно однакові. В той же час, для періоду 507 діб перевірка нульової гіпотези відкинула деякі з аномалій і в результаті це дозволило досягти кращої точності роботи моделі Prophet.

Отже, оптимальною моделлю є модель зазначена як Варіант 5 – модель для періоду 419 днів зі статистично значущими аномаліями. З порівняння точності оптимальної моделі з точністю моделі FB Prophet з параметрами і видами сезонності за замовчуванням (добова, тижнева та річна) видно, що оптимальна модель є значно точнішою, зокрема, за метрикою R2 (чим ближче до 1, тим – краще) – в 1,7 рази більше, а за метрикою MSE (чим ближче до нуля, тим – краще) – у 2 рази менше. Результат роботи оптимальної моделі наведено на рис. 3.13.

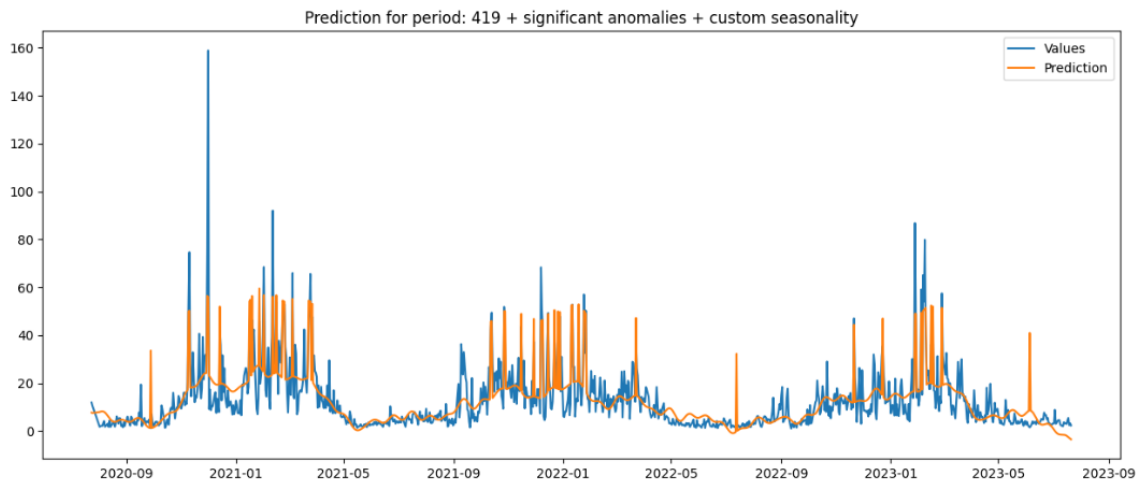


Рисунок 3.13 – Результат роботи оптимальної моделі FB Prophet, ідентифікованої за запропонованим методом [25, 43]

3.2. Метод ідентифікації локальних аномалій часового ряду

Для побудови інтелектуальної моделі на основі моделі Prophet потрібно, передусім, ідентифікувати аномалії. Як було згадано в пункті 1.3.4 та підрозділі 1.6, ці аномалії можуть бути локальними для певного інтервалу часового ряду.

Для розв’язання задачі пропонується метод, основна ідея якого це гіпотеза про те, що часовий ряд складається з окремих фрагментів між локальними мінімумами тренду – точками, де графік міняє свій напрям зі спадання на зростання, і на кожному з цих фрагментів ряд поводить себе по різному [44]. Відповідно, замість застосування певних методів пошуку аномалій до усього ряду, пропонується аналізувати кожен з фрагментів окремо, а потім об’єднувати усі отримані аномалії в один набір. Як було сказано раніше, показники стану довкілля мають циклічність, обумовлену впливом природних факторів. В якості прикладу подібного ряду, наведемо графік значень показнику пилу PM10 за 2020-2023 рр. у Вінницькій області. Графік зображено на рис. 3.14.

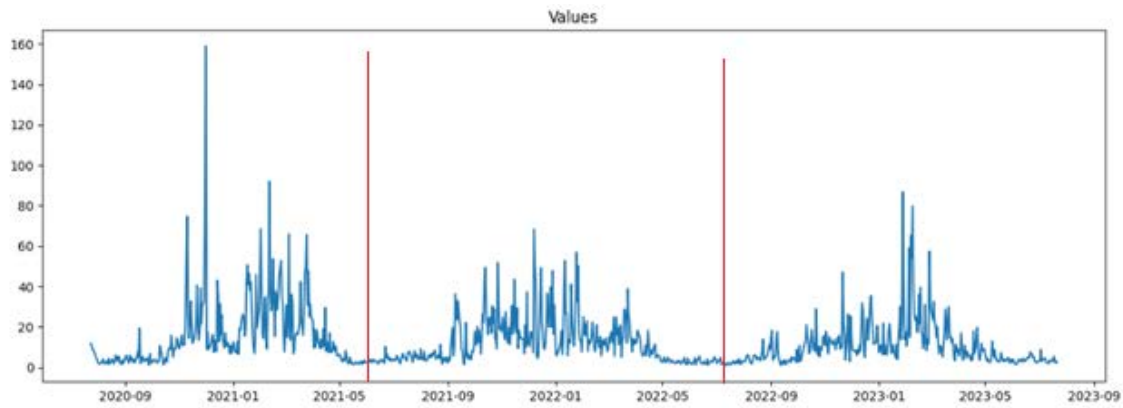


Рисунок 3.14 – Часовий ряд значень показника PM10 з приблизним поділом на окремі фрагменти [44]

Як схематично зображено на рис. 3.14, часовий ряд можна розділити на декілька фрагментів, які мають схожу поведінку, обумовлену певними природними факторами. Попередній досвід авторів показує, що розгляд усього ряду як одне ціле не завжди дає гарний результат і доцільним є пошук локальних аномалій в кожному окремо декомпозованому фрагменті ряду, а потім об'єднати результати по кожному фрагменту в одну вибірку.

Опишемо етапи роботи запропонованого методу.

Етап 1. Виділення тренду часового ряду.

Для виділення тренду скористаємося відомим прийомом – методом «rolling». Цей метод дозволяє згладити коливання даних та виокремити тренд часового ряду. Щоб обмежитися меншою кількістю сегментів та тим самим спробувати захопити більше подібних точок часового ряду, отриману криву тренду додатково згладжуємо певним алгоритмом згладжування даних. Результатом цього етапу для прикладу, наведеного на рис 3.14, буде графік зображений на рис. 3.15.

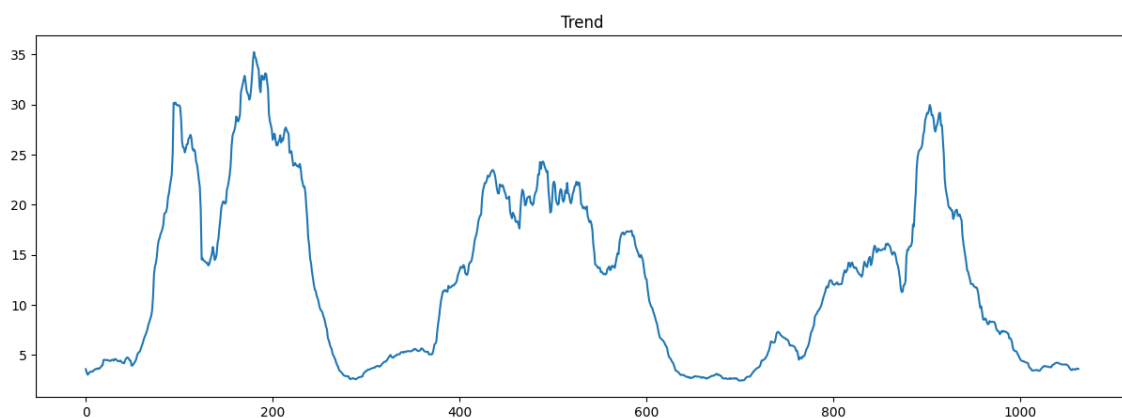


Рисунок 3.15 – Згладжена складова тренду для часового ряду показника PM10, наведеного на рис. 3.14 [44]

Етап 2. Пошук локальних мінімумів та максимумів та виділення окремих фрагментів ряду.

На отриманій згладженій кривій тренду часового ряду, необхідно знайти локальні мінімуми та локальні максимуми – точки кривої, де спадання міняється на зростання та навпаки. Крім того, варто знайти та ігнорувати ті локальні максимуми, де відстань до наступного локального максимуму є малою. Таким чином ми зможемо відкинути незначні коливання та не аналізувати невеликі фрагменти як окремі складові часового ряду, оскільки пошук аномалій на малих часових відрізках не є доцільним. Приклад локальних мінімумів та локальних максимумів для ряду, зображені на рис. 3.16.

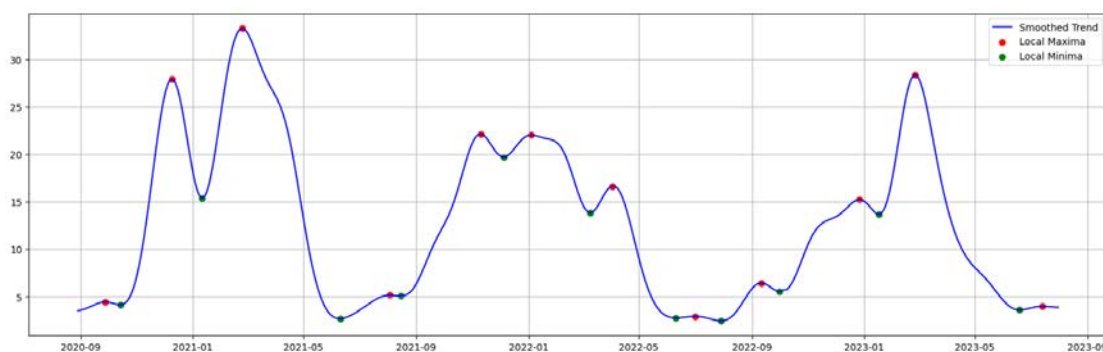


Рисунок 3.16 – Локальні мінімуми та локальні максимуми на кривій тренду часового ряду [44]

Виділивши з загального списку вагомі локальні мінімуми, ми отримуємо точки по яких розділяємо наш часовий ряд на декілька фрагментів. Кількість цих фрагментів залежить від самого часового ряд, параметрів згладжування та визначення вагомості. Останні два значення можуть бути винесені як параметри налаштування методу задля більш гнучких можливостей аналізу відповідного часового ряду.

Етап 3. Пошук аномалій на окремих сегментах часового ряду, використовуючи комбінації методів.

Як було описано в попередньому розділі, існує велика кількість різних за своєю природою методів для пошуку аномалій в часових рядах. З порівняльного аналізу видно, що кожен з методів має як свої переваги, так і свої недоліки. Як засіб усунення недоліків одного методу, пропонується одночасне використання декількох за різними стратегіями їх комбінування. Подібне поєднання може дозволити досягти більшої точності, оскільки один метод може знайти ті аномалії, які були пропущені іншим методом або ж навпаки, два метода будуть валідувати один одного і відкидати хибні аномалії.

Для цього пропонується три стратегії поєднання методів:

Варіант 1. Одночасне застосування двох методів та об'єднання їх

результатів. Нехай вхідний набір даних це X , тоді аномальні значення, отримані даним комбінованим методом можна позначити як

$$V_1(X) = M_1(X) \cup M_2(X), \quad (3.7)$$

де $M_1(X)$ – аномальні дані, отримані першим методом для даних X , а $M_2(X)$ – аномальні дані, отримані другим методом.

Варіант 2. Одночасне застосування двох методів та знаходження перетину їх результатів. Нехай вхідний набір даних це X , тоді аномальні значення, отримані даним комбінованим методом можна позначити як

$$V_2(X) = M_1(X) \cap M_2(X), \quad (3.8)$$

де $M_1(X)$ – аномальні дані, отримані першим методом для даних X , а $M_2(X)$ – аномальні дані, отримані другим методом.

Варіант 3. Покрокове застосування обох методів, де після застосування першого методу, з набору вхідних даних видаляються аномальні дані і до отриманого набору застосовується другий метод. Результат роботи обох методів об'єднується в один набір. Нехай вхідний набір даних це X , а A_1 – аномальні значення отримані за допомогою першого методу пошуку аномалій, тоді аномальні значення, отримані даним комбінованим методом можна позначити як

$$V_3(X) = A_1 \cup M_2(X \setminus A_1). \quad (3.9)$$

Усі можливі комбінації методів пошуку аномалій потрібно застосовувати на кожному фрагменті ряду, виділених на попередньому кроці. Таким чином, слідуючи нашій гіпотезі, що часовий ряд показників стану довкілля поводить себе по різному на кожному окремому відрізку ряду, ми зможемо проаналізувати фрагмент окремо від інших частин часового ряду та ідентифікувати локальні аномалії, властиві безпосередньо цьому відповідному фрагменту ряду.

Блок-схема алгоритму запропонованого методу наведена на рис. 3.17.

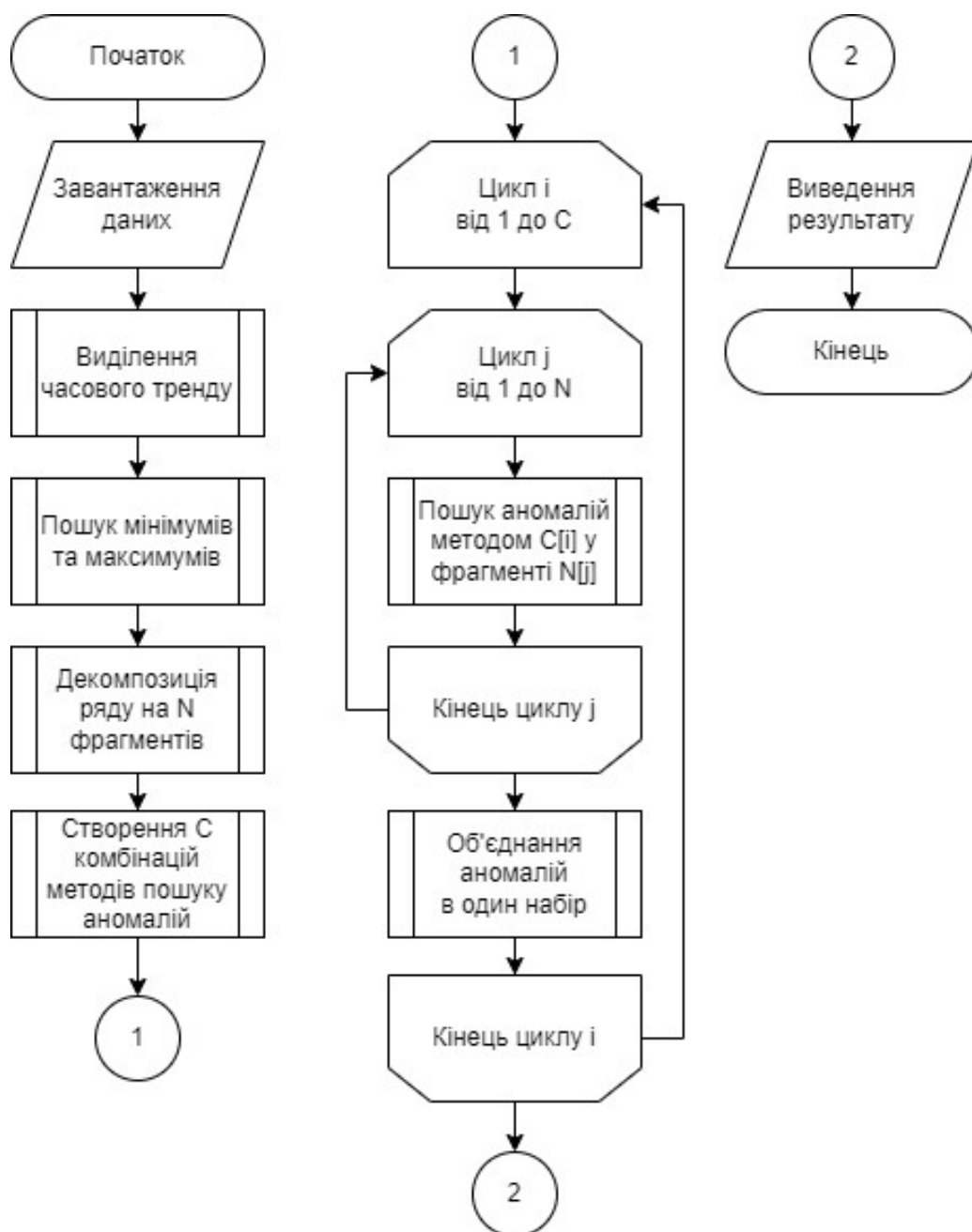


Рисунок 3.17 – Блок-схема алгоритму запропонованого методу ідентифікації локальних аномалій [44]

Для перевірки методу знову використаємо дані показників якості атмосферного повітря від мережу громадського моніторингу EcoCity. В цьому конкретному випадку було обрано одну зі станцій, розташованих безпосередньо у місті Вінниця та показник пилу PM2.5 (мікроскопічні тверді частинки) за період 2019-2022 рр. Графік цих даних наведено на рис. 3.18.

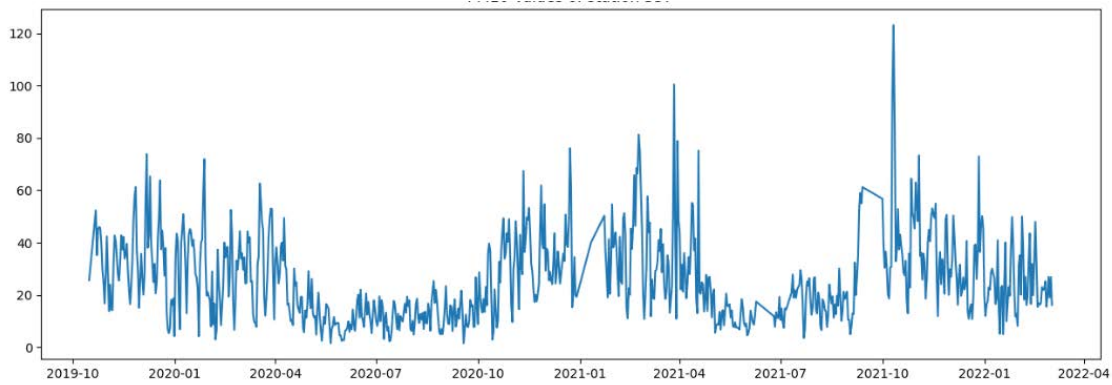


Рисунок 3.18 – Дані показника PM2.5 станції у місті Вінниця [44]

Покроково застосуємо метод ідентифікації.

Етап 1. Знайдемо складову тренду для даних, зображених на рис. 3.17 та згладимо їх за допомогою фільтру Баттерворта [40] – одного з модулів Python-бібліотеки `scipy.signal`. Крива тренду та її згладжений варіант наведено на рис. 3.19.

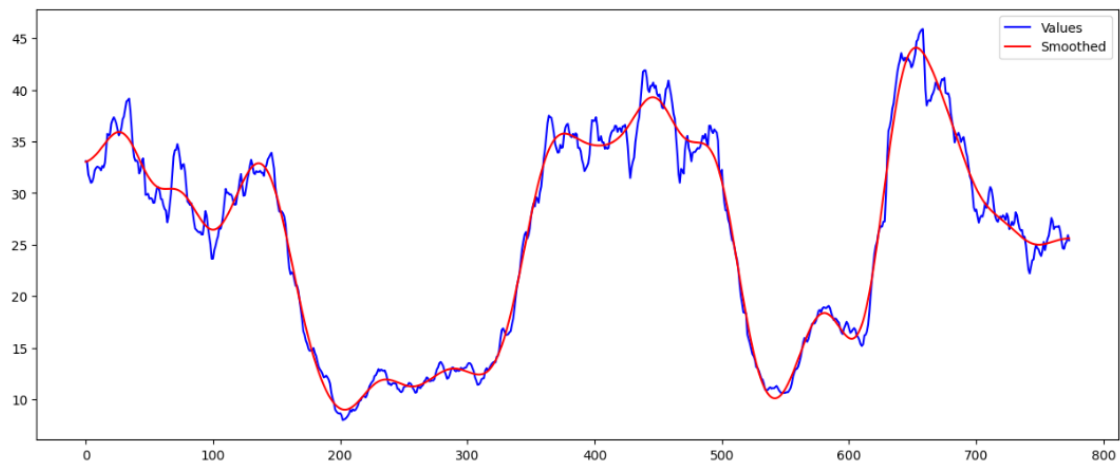


Рисунок 3.19 – Складова тренду та її згладжена версія для часового ряду, зображеного на рис. 3.18 [44]

Етап 2. Визначимо локальні мінімуми та локальні максимуми на згладженій кривій тренду. Для цього застосуємо методи з раніше згаданої Python-бібліотеки `scipy.signal`. Отримані значення зображено на рис. 3.20.

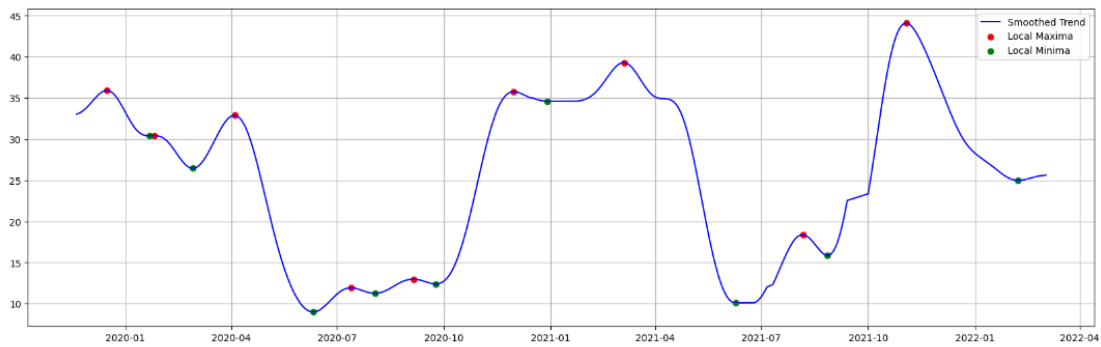


Рисунок 3.20 – Локальні максимуми та локальні мінімуми кривої тренду [44]

Порівнюючи сусідні мінімуми та максимуми, необхідно знайти ті, які не суттєво відрізняються один від одного. Такі точки локальних мінімумів, як було описано в попередньому розділі будуть ігноруватися та не будуть використовуватися для визначення сегментів часового ряду [40]. Важливі локальні мінімуми зображено на рис. 3.21.

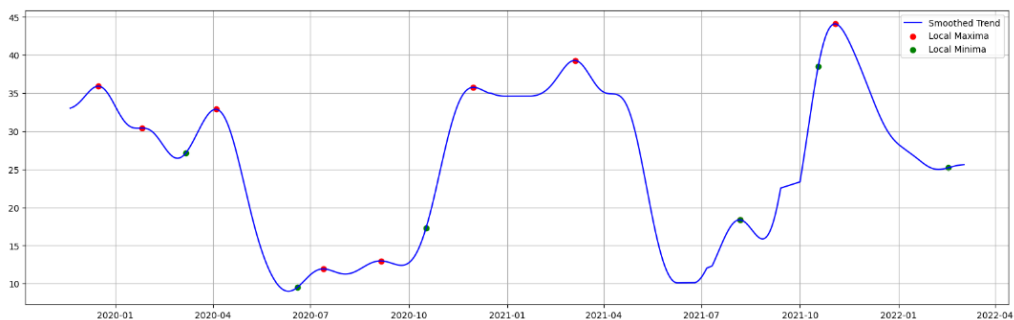


Рисунок 3.21 – Значущі локальні мінімуми кривої тренду [44]

Розділимо часовий ряд на окремі сегменти, використовуючи отримані точки значущих локальних мінімумів. Результат зображено на рис. 3.22.

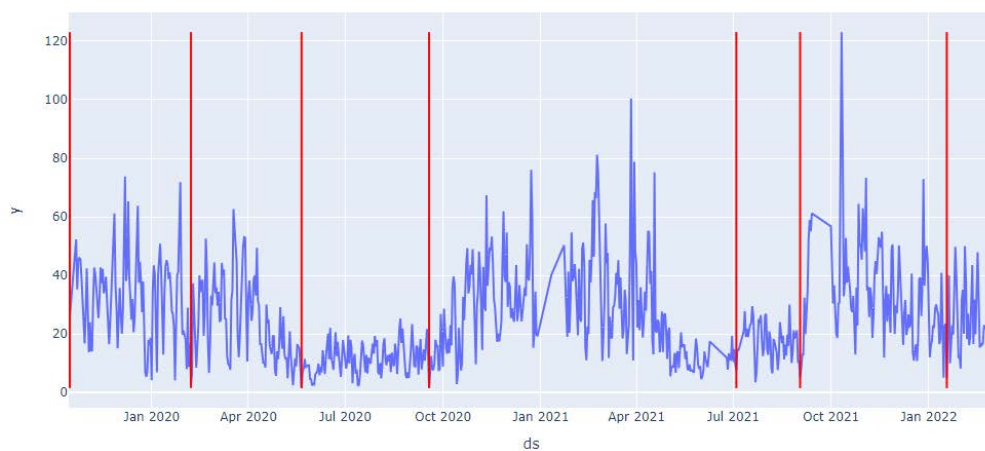


Рисунок 3.22 – Окремі сегменти часового ряду розділені червоними лініями – точками локальних мінімумів [44]

Етап 3. Виконаємо пошук аномалій на кожному окремому фрагменті часового ряду, використовуючи наш комбінований підхід. Для цього використаємо 7 методів пошуку аномалій, описаних раніше в порівняльному аналізі. Крім використання різних комбінацій цих методів, варто також спробувати застосувати кожен з них окремо і тим самим порівняти чи дає комбінування методів якусь перевагу у точності порівнюючи з окремими методами пошуку аномалій. Як було описано, комбінування методів буде виконуватися за трьома стратегіями. Перші дві стратегії комбінування не залежать від порядку використання методу пошуку, що відповідає вибірці, яка називається комбінація без повторень:

$$N_1 = \frac{n!}{r!(n-r)!}, \quad (3.10)$$

де n – загальна кількість методів, r – кількість методів, які застосовуються разом, а N_1 – загальна кількість можливих унікальних комбінацій.

В нашому випадку усього є 7 методів пошуку аномалій і кожна комбінація включає 2 методи при застосуванні першої стратегії комбінування методів пошуку аномалій. Підставивши ці числа в формулу 3.10, ми отримуємо 21 можливу комбінацію. Таке саме число комбінацій є і при застосуванні другої стратегії комбінування методів пошуку аномалій.

В той же час, при застосуванні третьої стратегії комбінування методів пошуку аномалій, результат буде різним в залежності від порядку застосування методу. Таким чином, загальну кількість комбінацій можна знайти за формулою:

$$N_2 = n \times (n-1), \quad (3.11)$$

де n – загальна кількість методів, а N_2 – загальна кількість комбінацій з можливими перестановками.

Оскільки використовуються 7 різних методів пошуку аномалій, то загальна кількість комбінацій згідно формули 3.11 дорівнює 42. Крім того, також застосовуються кожен з семи методів окремо, що додає ще 7 варіантів пошуку аномалій. Таким чином, загалом приклад має 91 варіант. Реалізуючи цей перебір за допомогою Python, отримуємо результат, зображений на рис. 3.23.

	name	anomaly_values
0	Hampel	[13, 49, 75, 76, 93, 130, 170, 215, 239, 260, ...
1	Moving Average	[46, 100, 110, 125, 148, 152, 153, 161, 260, 2...
2	Isolation Forest	[46, 98, 148, 149, 305, 316, 332, 465, 497, 59...
3	LOF	[1, 34, 35, 36, 45, 46, 49, 57, 58, 66, 71, 79...
4	k-neighbors	[1, 34, 35, 36, 45, 46, 49, 58, 79, 97, 98, 10...
...	...	...
86	SVM → Moving Average anomalies	[1, 6, 28, 36, 45, 46, 47, 49, 50, 54, 61, 63,...
87	SVM → Isolation Forest anomalies	[1, 6, 28, 45, 46, 47, 49, 50, 61, 63, 71, 77,...
88	SVM → LOF anomalies	[1, 6, 28, 32, 33, 45, 46, 47, 49, 50, 56, 61,...
89	SVM → k-neighbors anomalies	[1, 6, 27, 28, 31, 32, 33, 45, 46, 47, 49, 50,...
90	SVM → SESD anomalies	[1, 6, 13, 23, 27, 28, 32, 33, 42, 45, 46, 47,...

Рисунок 3.23 – Результат пошуку аномалій кожним з 91-го варіанту перебору [44]

Для перевірки точності роботи експертним шляхом знайдемо аномальні записи на кожному з фрагментів часового ряду. В ході експертного аналізу було виділено 41 запис, який візуально відрізняється від інших записів відповідного фрагменту часового ряду. Перелік аномалій, знайдених експертним шляхом, зображено на рис. 3.24.

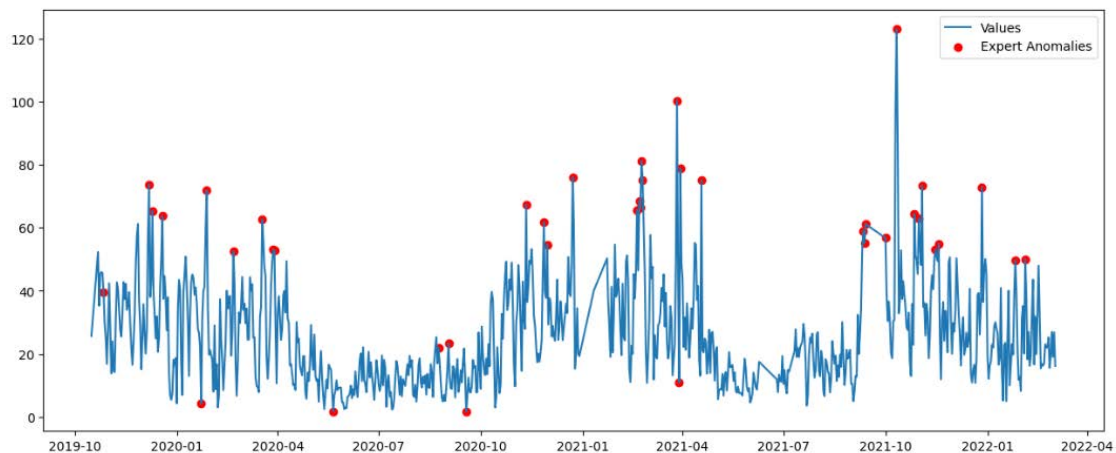


Рисунок 3.24 – Аномалії часового ряду, визначені експертним шляхом [44]

Для перевірки точності кожної з 91 комбінації методів пошуку аномалій, перевіримо наскільки список експертних аномалій відрізняється від відповідного списку аномалій, знайдених програмним шляхом. Для цього створимо список значень, який дорівнює розміру нашого часового ряду. Кожен елемент списку це 1, якщо відповідний елемент ряду вважається аномалією або 0, якщо елемент вважається но-

рмальним. Як критерій точності використаємо декілька статистичні метрики: влучність (англ. «precision» або «precision score»), повнота (англ. «recall» або «recall score») та f-міра – показник, який обчислюється через значення повноти та влучності [28]. Обрахувавши ці метрики для кожного з списків, ми отримуємо результат як це показано на рис. 3.25.

	name	anomaly_values	precision_score	recall_score	f1_score
43	LOF $\cap$ k-neighbors anomalies	[1, 34, 35, 36, 45, 46, 49, 58, 79, 97, 98, 10...	0.737571	0.830996	0.775075
18	Isolation Forest + LOF anomalies	[1, 34, 35, 36, 45, 46, 49, 57, 58, 66, 71, 79...	0.705752	0.836943	0.751830
3	LOF	[1, 34, 35, 36, 45, 46, 49, 57, 58, 66, 71, 79...	0.705752	0.836943	0.751830
69	LOF $\rightarrow$ Isolation Forest anomalies	[1, 34, 35, 36, 45, 46, 49, 57, 58, 66, 71, 79...	0.702652	0.836288	0.749021
45	LOF $\cap$ SVM anomalies	[1, 45, 46, 49, 71, 93, 98, 108, 148, 149, 158...	0.737516	0.748739	0.742974
4	k-neighbors	[1, 34, 35, 36, 45, 46, 49, 58, 79, 97, 98, 10...	0.687918	0.844201	0.737270
19	Isolation Forest + k-neighbors anomalies	[1, 34, 35, 36, 45, 46, 49, 58, 79, 97, 98, 10...	0.687918	0.844201	0.737270
47	k-neighbors $\cap$ SVM anomalies	[1, 45, 46, 49, 98, 108, 148, 149, 158, 159, 2...	0.736894	0.736894	0.736894
75	k-neighbors $\rightarrow$ Isolation Forest anomalies	[1, 34, 35, 36, 45, 46, 49, 50, 58, 79, 97, 98...	0.685411	0.843545	0.734748
44	LOF $\cap$ SESD anomalies	[34, 35, 36, 45, 46, 49, 57, 58, 66, 71, 79, 9...	0.750709	0.702670	0.723725
63	Isolation Forest $\rightarrow$ LOF anomalies	[1, 34, 35, 36, 45, 46, 48, 56, 57, 65, 70, 78...	0.681653	0.797477	0.722049
22	LOF + k-neighbors anomalies	[1, 34, 35, 36, 45, 46, 49, 57, 58, 66, 71, 79...	0.669843	0.850147	0.719860
64	Isolation Forest $\rightarrow$ k-neighbors anomalies	[1, 34, 35, 36, 45, 46, 48, 56, 57, 78, 92, 96...	0.668928	0.816579	0.714111
14	Moving Average + LOF anomalies	[1, 34, 35, 36, 45, 46, 49, 57, 58, 66, 71, 79...	0.665431	0.837647	0.713570
76	k-neighbors $\rightarrow$ LOF anomalies	[1, 34, 35, 36, 45, 46, 49, 50, 58, 63, 79, 84...	0.665431	0.837647	0.713570
46	k-neighbors $\cap$ SESD anomalies	[34, 35, 36, 45, 46, 49, 58, 79, 148, 149, 237...	0.742520	0.690170	0.712649

Рисунок 3.25 – Результати обрахунку точності кожного з методів пошуку аномалій у порівнянні з аномаліями, отриманими експертним шляхом, посортовані за показником f1_score [44]

Як можна помітити, серед найкращих методів часто зустрічається метод LOF. Також високої точності мають методи з використанням Isolation Forest та k-neighbors. Якщо намагатися роздивлятися які стратегії комбінування методів були найефективнішими, то видно що серед найкращих зустрічаються усі три підходи.

Як додатковий етап перевірки, спробуємо використати найточніший метод для апроксимації часового ряду. Для цього побудуємо модель Prophet з параметрами за замовчування та з налаштуванням «holidays», який дозволяє передати в модель список аномальних дат і тим самим покращити роботу моделі. Результати роботи моделі Prophet з параметрами за замовчуванням та з використанням аномалій зображено на рис. 3.26 та рис. 3.27 відповідно.

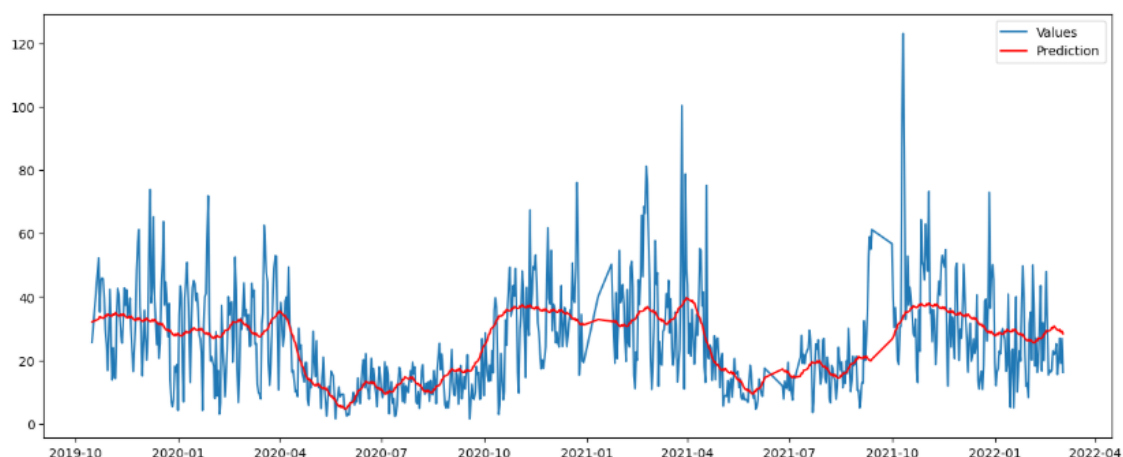


Рисунок 3.26 – Результати роботи моделі Prophet з параметрами налаштування за замовчуванням [44]

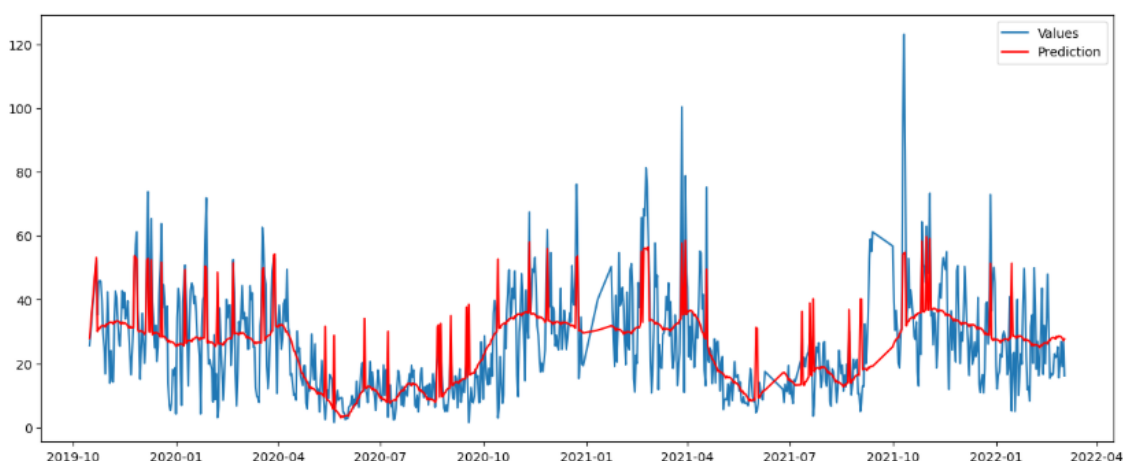


Рисунок 3.27 – Результати роботи моделі Prophet та аномалій, отриманих найточнішим комбінованим методом [44]

Точність апроксимації було обраховано за допомогою середньої абсолютної похибки (англ. «Mean absolute error» або скорочено – MAE) та середньоквадратична похибка (англ. «Root mean squared error» або скорочено – RMSE) [28]. За цим показником похибка апроксимації зменшилась на 11% за RMSE та 8% за MAE при використанні ідентифікованих аномалій у порівнянні з моделлю Prophet з параметрами за замовчуванням. Варто зазначати, що обраний часовий ряд досить важко аналізувати та передбачувати, але все рівно вдалось підняти точність роботи моделі.

3.3. Пошук взаємовпливів показників стану атмосферного повітря на основі байєсівського моделювання

Однією з проблем аналізу якості атмосферного повітря та пошуку аномального впливу на його стан є складність механізмів розповсюдження забруднювальних речовин, які можуть залежати від різних природних та географічних умов. Більш традиційні методи моніторингу та аналізу дозволяють отримати поверхневу картину загального стану, проте не завжди дають чітке розуміння взаємозв'язків між показниками забруднення, отриманих з різних точок.

Наявність певного аномального показника забруднення атмосферного повітря в певній географічній точці зазвичай повинно впливати та проявлятися в інших місцях, наприклад, підвищення рівня пилу на одній з станцій моніторингу може передувати подібному підвищенню рівня забруднення на іншій станції за рахунок напрямку вітру, який переносить даний пил.

Одним із можливих підходів для аналізу залежності між різними чинниками забруднення або між різними станціями моніторингу цього забруднення є байєсівські мережі, які дозволяють ідентифікувати причинно-наслідкові моделі на основі існуючих даних. Даний метод дозволяє яким чином зміни значень показника впливають на зміну цього або іншого показника в межах певної географічної зони, що, в свою чергу, може допомогти локалізувати джерело забруднення та вживати відповідні заходи для зниження ризиків для здоров'я населення та захисту довкілля.

В одній з попередніх робіт двох авторів даної монографії [45] було застосовано байєсівський аналіз для моделювання чутливості людей до певних грибкових алергенів. Байєсівська мережа застосовувалась для визначення ймовірності розвитку комбінованої чутливості в окремих пацієнтів до молекулярних компонентів грибів. Даний метод можна адаптувати та застосувати для аналізу взаємовпливу показників стану атмосферного повітря, а саме в моделюванні ймовірностей зміни значення показника якості атмосферного повітря за умов зміни значень відповідного показника на сусідніх станціях мережі моніторингу.

Алгоритм даного підходу можна описати наступним чином:

1. *Підготовка даних.* Дані моніторингу якості атмосферного повітря зазвичай включають абсолютні значення відповідних показників в часі. Загалом для байєсівського моделювання є можливим використання безпосередньо абсолютні значення, проте в класичному варіанті все ж варто перетворити ці абсолютні значення у бінарний формат. В нашому випадку пропонується орієнтуватися на зміну відповідного показника якості

повітря. Запропонований підхід до підготовки даних полягає в порівняння значення показника з його попереднім значенням. Якщо поточний показник є більшим за попередній, то його значення замінюється на «1», якщо ж значення менше або рівне, то значення замінюється на «0».

Ці кроки необхідно провести з значеннями відповідного показника, отриманих з інших станцій моніторингу та зібрати їх в один набір та виділити з нього тренувальні дані.

2. *Побудова спрямованого ациклічного графа.* Для отриманих даних ми проводимо вивчення їх структури. Це дозволяє виявити залежності між наявними змінними та згенерувати спрямований ациклічний граф.

3. *Вивчення параметрів.* Вивчення параметрів даних – це оцінка значень умовних розподілів ймовірностей (CPD). Для цього використовується спрямований ациклічний граф, отриманий на попередньому етапі. Результатом даного етапу є набір імовірностей зміни значення певної змінної графу в залежності від зміни відповідних батьківських вершин графу.

Для перевірки роботи цього методу було вирішено застосувати його для даних моніторингу якості атмосферного повітря мережі громадського моніторингу EcoCity з декількох станцій моніторингу у Вінницькій області за період 2022-2024 рр. Для тестування буде використовуватися показник пилу PM10. Приклад таких даних наведено на рисунку 3.28.

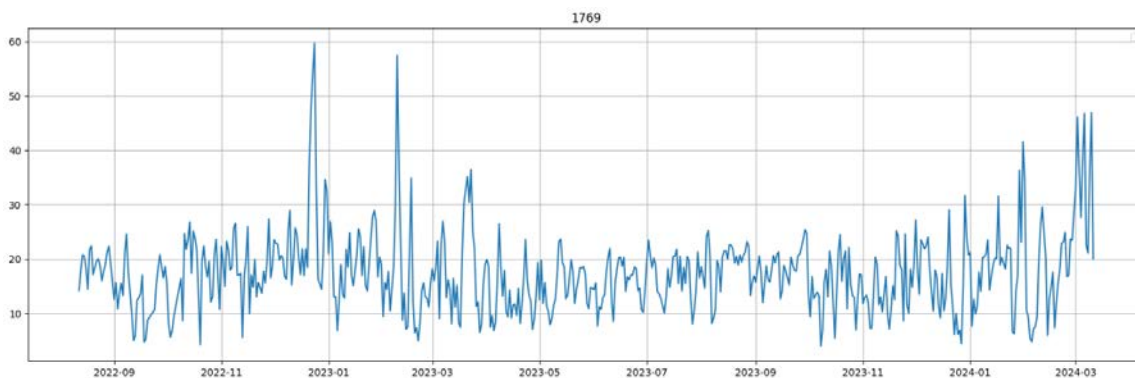


Рисунок 3.28 – Дані показника якості атмосферного повітря однієї з станцій моніторингу у Вінницькій області за 2022-2024 рр [46]

Як було описано раніше, для застосування методу необхідно зібрати дані з декількох станцій та створити один набір, де кожен рядок це дата проведення вимірювання, а значення в колонках це значення показника PM10 для відповідної станції моніторингу. Загалом було побудовано набір даних з семи станцій моніторингу (рис 3.29).

	st_650	st_1769	st_1811	st_1813	st_1825	st_1830	st_1834
0	9.4885	7.4381	11.8504	8.0747	12.7674	9.7900	9.9807
1	31.2040	8.3401	10.1477	8.9014	13.2750	22.9919	10.7449
2	11.8139	14.1653	13.9522	22.2361	18.0257	10.8561	19.7283
3	8.3002	15.6867	19.0788	27.1034	27.0721	18.1392	20.3049
4	14.4049	13.0627	23.9908	15.5545	22.7843	12.5923	15.0530
...	...	...	...	...	...	...	...
279	48.6737	22.5754	39.2001	30.8495	30.3635	29.2202	33.4889
280	17.6561	21.1076	47.1072	32.4196	39.6464	43.1531	32.4889
281	26.1567	36.6581	68.6969	68.4406	74.4583	38.1823	36.8276
282	19.1412	46.9558	56.4339	65.3547	51.9435	91.6873	52.6314
283	7.0339	20.1094	15.9684	16.2481	17.2579	12.3558	13.9910

Рисунок 3.29 –Набір даних для байєсівського аналізу [46]

Наступним кроком є перетворення отриманих даних у бінарний формат, використовуючи описаний раніше підхід, де кожне значення для відповідної станції порівнюється з попереднім спостереженням для визначення чи рівень забруднення зростає, чи спадає (рис 3.30).

	st_650	st_1769	st_1811	st_1813	st_1825	st_1830	st_1834
0	0	0	0	0	0	0	0
1	1	1	0	1	1	1	1
2	0	1	1	1	1	0	1
3	0	1	1	1	1	1	1
4	1	0	1	0	0	0	0
...	...	...	...	...	...	...	...
279	1	0	0	0	0	0	0
280	0	0	1	1	1	1	0
281	1	1	1	1	1	0	1
282	0	1	0	0	0	1	1
283	0	0	0	0	0	0	0

Рисунок 3.30 – Дані після перетворення, згідно кроку 2 алгоритму [46]

Для отриманого набору даних побудуємо спрямований ациклічний граф (рис 3.31), використовуючи Python бібліотеку bnlearn. Дана бібліотека дозволяє моделювати байєсівські мережі та використовувати широкий набір допоміжних методів для малювання графіків.

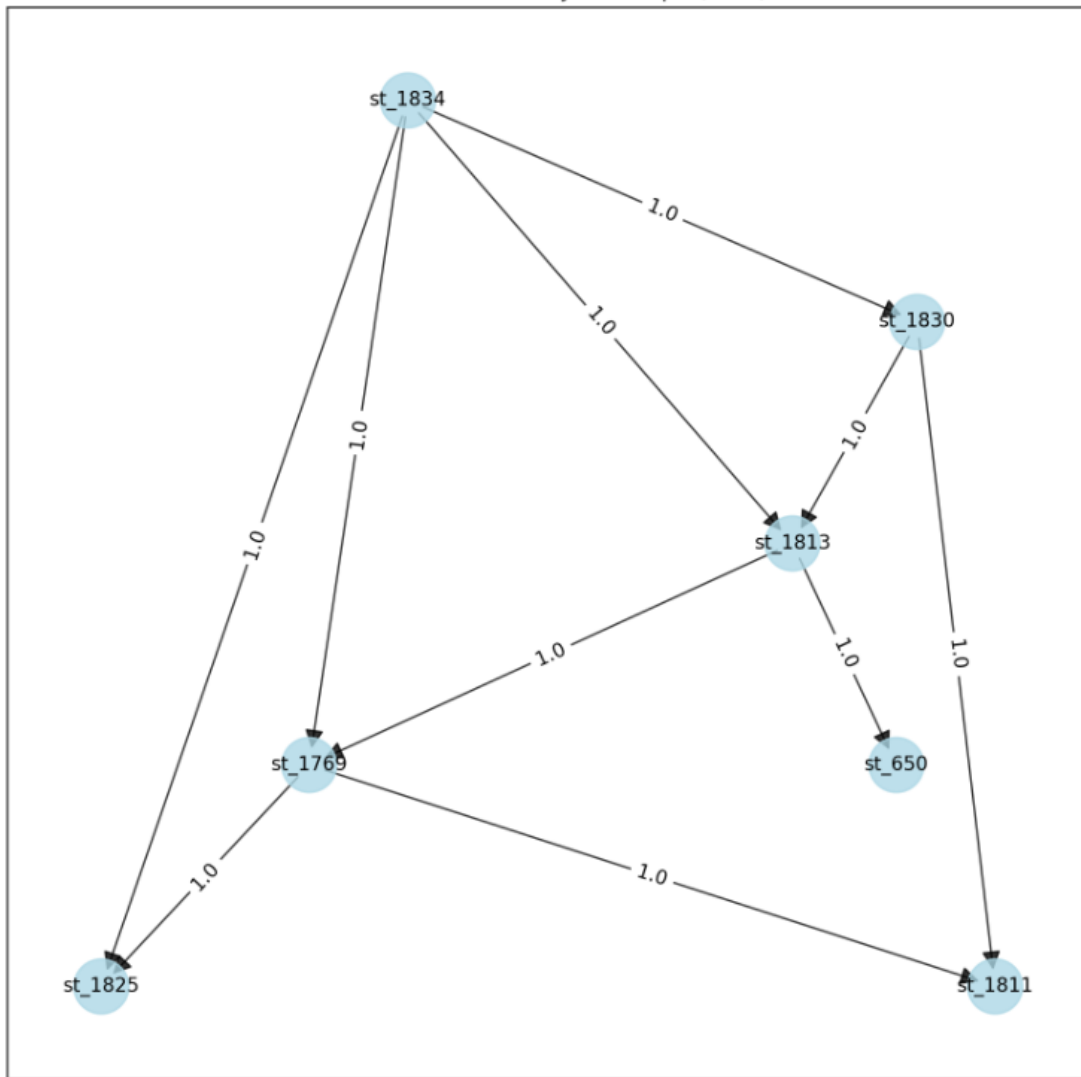


Рисунок 3.31 – Результат роботи бібліотеки bnlearn [46]

Як можна побачити, використовуючи бібліотеку для отриманих раніше даних вдалось встановити зв'язки між різними станціями моніторингу та побудувати DAG. Використаємо його для оцінки значення умовних розподілів ймовірностей (CPD). Для цього можна застосувати вже згадану бібліотеку bnlearn, використовуючи модуль «parameter_learning» Результат роботи даного модулю наведено на рисунках 3.32 та 3.33.

```
[bnlearn] >CPD of st_1769:
```

+-----+-----+-----+-----+			
st_1813	st_1813(0)	... st_1813(1)	
+-----+-----+-----+-----+			
st_1834	st_1834(0)	... st_1834(1)	
+-----+-----+-----+-----+			
st_1769(0)	0.5813253012048193	... 0.3994252873563218	
+-----+-----+-----+-----+			
st_1769(1)	0.4186746987951807	... 0.6005747126436781	
+-----+-----+-----+-----+			

Рисунок 3.32 – Оцінка значення умовних розподілів ймовірностей для станції 1769 [46]

```
[bnlearn] >CPD of st_1811:
```

+-----+-----+-----+-----+			
st_1769	st_1769(0)	... st_1769(1)	
+-----+-----+-----+-----+			
st_1830	st_1830(0)	... st_1830(1)	
+-----+-----+-----+-----+			
st_1811(0)	0.5772870662460567	... 0.4057971014492754	
+-----+-----+-----+-----+			
st_1811(1)	0.4227129337539432	... 0.5942028985507246	
+-----+-----+-----+-----+			

Рисунок 3.33 – Оцінка значення умовних розподілів ймовірностей для станції 1811 [46]

Як можна побачити на рисунку 3.29, за умови зростання забруднення на станціях «1813» (с. Якушинці) та «1834» (с. Стрижавка), імовірність зростання забруднення на станції «1769» (м. Вінниця) зростає в 1.5 рази. Подібний результат можна побачити і на рисунку 3.30, а саме імовірність росту забруднення на станції «1811» (м. Хмільник) зростає майже в 1.5 рази за умови росту забруднення на станціях «1769» (м. Вінниця) та «1830» (с. Війтівці).

Під час перевірки точності роботи моделі було отримано значення точності 77% (рис 3.34).

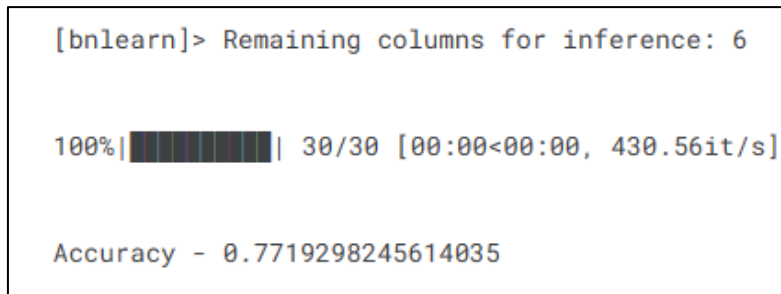


Рисунок 3.34 – Точність роботи моделі [46]

3.4. Алгоритм системного аналізу стану атмосферного повітря регіону

Узагальнюючи описані в попередніх пунктах і підпунктах методи розкриття невизначеності, запропонуємо узагальнений алгоритм системного аналізу стану атмосферного повітря регіону:

1. Вибір критеріїв стану атмосферного повітря: обрати критерій та його складові, за якими буде аналізуватися цей стан, що розкриває невизначеність U_{T2} .

2. Збирання інформації відповідно до моделі «DPSIR»: зробити запити для отримання інформації про місця розташування джерел забруднення (для розкриття невизначеності U_{I1}), про розташування джерел викидів чи скидань та фактичні обсяги і концентрації забруднювальних речовин у цих викидах (скиданнях) (U_{I3}), про найбільші обсяги гранично допустимих значень цих обсягів та концентрацій (U_{I2}); щодо надання рядів даних спостережень за певний період часу (U_{I4}); про вже прийняті управлінські рішення щодо удосконалення моніторингу, контролю та охорони стану довкілля, боротьби з рівнем забруднення (U_{I6}).

3. Аналіз даних спостережень про «S»: проаналізувати *чи наявних даних про стан атмосферного повітря достатньо* для формування висновків та управлінських заходів, використовуючи прийоми розвідувального аналізу. Якщо – так, тоді слід перейти до кроку 7, а якщо – ні, тоді – слід оптимізувати систему моніторингу.

4. Збирання інформації про мережу моніторингу та контролю «SIR». Слід зібрати інформацію від суб'єктів державної системи моніторингу стану атмосферного повітря про пости цієї мережі, їх показники, регулярність вимірювань та метрологічні аспекти (це розкриє U_{I5}) та зібрати додаткову інформацію про стан атмосферного повітря від суб'єктів державної системи контролю, які здійснюють нерегулярні спостереження (держкоінспекції, органи МОЗ тощо), від мереж громадського моніторингу, непрямі дані про щільність населення, зокрема

вразливого, про захворюваність людей в регіоні тощо.

5. Аналіз аномалій у «S»: відповідно до чинного законодавства (наприклад, для атмосферного повітря – за Постановою КМУ від 14.08.2019 р. № 827) або за результатами статистичного аналізу, визначити верхнє і нижнє порогові значення відповідних показників та кількість фактів виходу за межі цих порогових значень, які потім будуть застосовуватися для пошуку аномалій чи екстремальних значень показників стану природного середовища – це розкриває невизначеності U_{T3} та U_{T5} . А потім на основі аналізу кількості таких аномалій за цією ж Постановою визначається мінімальна кількість постів мережі моніторингу для кожної групи показників окремо, що дозволяє розкрити невизначеність U_{T1} в частині кількості постів.

6. Оптимізація місць розташування постів мережі моніторингу «S». Сформувати систему критеріїв для вибору параметрів місць постів спостережень стану атмосферного повітря, з урахуванням екстремальних значень показників про усіх складові «DPSIR». Для їх систематизації та визначення ваги використати метод аналізу ієрархій, наприклад як це було зроблено у роботі для стану атмосферного повітря у Вінницькій області. Це дозволить визначити де варто розмістити нові станції моніторингу чи перемістити існуючі, що розкриє невизначеність U_{T1} в частині місць розташування цих станцій чи постів.

7. Проаналізувати достовірність даних про «Р». Якщо аналіз покаже певну невідповідність, наприклад, антропогенні впливи (наприклад, скидання забруднень у річку) не змінюються, а якість води біля місця впливу неухильно погіршується, тоді варто висунути та перевірити гіпотезу щодо недостовірності або даних про стан природного середовища, або даних про антропогенні впливи. Якщо немає підстав вважати дані недостовірними, тоді – перейти до кроку 9.

8. Пошук можливих джерел забруднення. Для цього можна застосувати авторський метод на основі аналізу аномалій у наявних даних моніторингу стану природного середовища. Це дозволить розкрити невизначеність (U_{S1}). Цей метод дозволяє знайти незареєстровані джерела (чи виявити який обсяг було занижено існуючим джерелом), які можуть мати місце, і в просторі, і в часі одночасно, шляхом зіставлення результатів роботи методу з наявною інформацією про складові моделі «DPSIR» заданого регіону. А це, у свою чергу, дозволить ще зменшити рівень невизначеності для U_{I3} .

9. Чи є необхідним прогнозування щодо «S»? Якщо – ні, тоді – перейти до кроку 12.

10. Визначення параметрів аномалій у «S». Для побудови ефе-

ктивної моделі прогнозування, необхідно знайти аномалії в наявних даних стану природного середовища. Також, дані можуть містити локальні у часі аномалії, тому можна застосувати метод пошуку локальних аномалій на основі декомпозиції ряду з роботи, що, в свою чергу, дозволить розкрити невизначеність U_{S2} .

11. Побудова оптимальної моделі для прогнозування «S». Це передбачає вибір можливих інтелектуальних моделей, ідентифікацію усіх їх параметрів, з урахуванням аномалій та можливих видів сезонності, їх порівняльний аналіз та вибір оптимальної за певною метрикою моделі. Це розкриває невизначеність U_{M1} . При цьому, для ідентифікації параметрів сезонності можна використати як класичні підходи, так і авторський метод ідентифікації параметрів сезонності та аномалій періодичного часового ряду, описаний у роботі, що дозволить знайти які сезонності мають місце для відповідного набору даних та яким чином ці сезонності можна поєднати при моделюванні даних. Усе це дозволить розкрити такі невизначеності: U_{P1} , U_{P2} , U_{T6} .

12. Чи є необхідним аналіз взаємозв'язків? Якщо – ні, тоді – перейти до кроку 14.

13. Інтелектуальний аналіз взаємозв'язків. Застосувати метод аналізу взаємовпливу показників стану природного середовища на основі байєсівського моделювання, що дозволить, хоча б частково, розкрити невизначеності U_{ST1} та U_{ST2} . В результаті можна буде отримати інформацію про те, який взаємовплив існує між різними показниками на сусідніх станціях моніторингу, що, у свою чергу, дозволить компенсувати проблеми з даними на різних станціях (наприклад, наявність пропусків через відсутність в певні проміжки часу зв'язку або електроенергії).

14. Аналіз даних «S», у т.ч. інтелектуальний, за різними сценаріями, що дозволить розкрити U_{SC1} .

15. Оптимізація складної системи «DPSIR». Виходячи з усіх отриманих результатів, провести оптимізацію різних складових моделі «DPSIR» та розробити рекомендації щодо можливих управлінських заходів системи і вибрати оптимальні серед них одним із відомих методів, що розкриє невизначеність U_{T4} .

Графічно даний алгоритм можна зобразити блок-схемою (рис. 3.35).

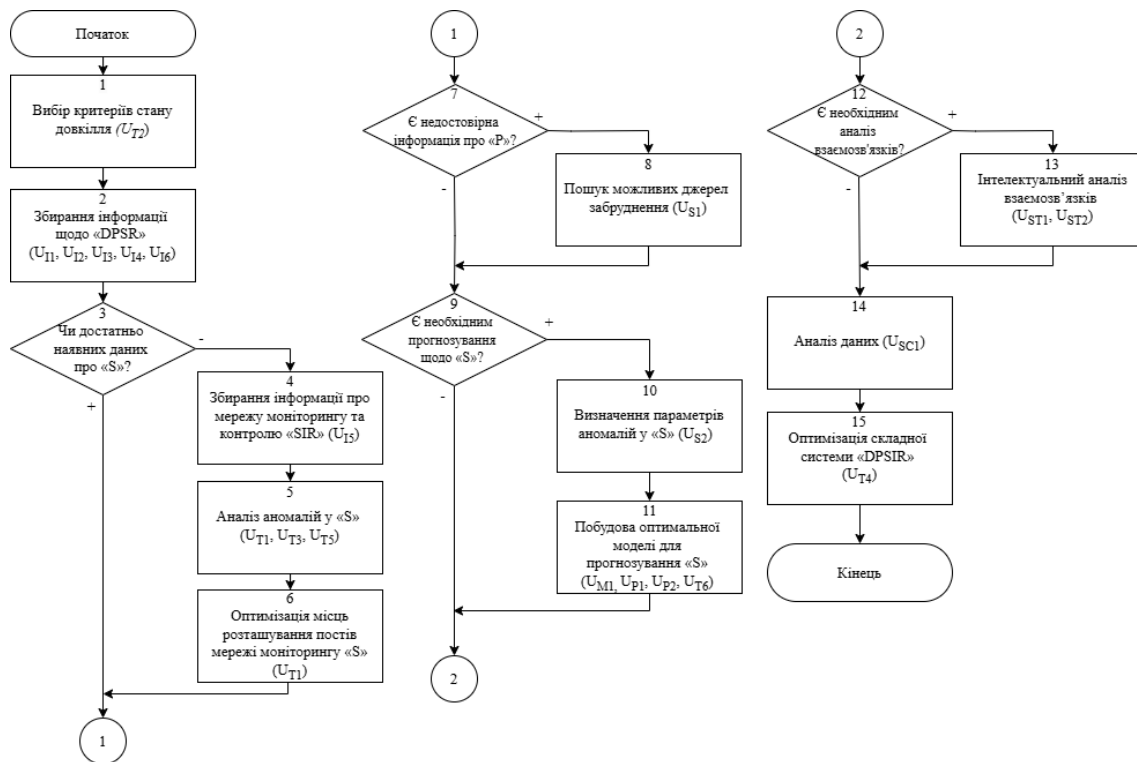


Рисунок 3.35 – Блок-схема алгоритму системного аналізу атмосферного повітря з розкриттям невизначеності

3.5. Висновки до розділу 3

Описано метод ідентифікації параметрів сезонності та аномалій періодичного часового ряду, що базується на адаптивній декомпозиції та дозволяє комбінувати налаштування декількох варіантів сезонності при роботі з інтелектуальними моделями прогнозування стану атмосферного повітря. Застосування методу дозволяє підвищити точність роботи моделі Prophet у 1.5-2 рази за різними метриками.

Описано метод ідентифікації локальних аномалій часового ряду на основі декомпозиції на півхвилі, що дозволяє шукати аномальні значення на різних інтервалах цього ряду та підвищувати точність прогнозування окремих відрізків часового ряду стану атмосферного повітря.

Описано алгоритм аналізу взаємозв'язків показників стану природного середовища, що базується на використанні байєсівського аналізу даних та дозволяє аналізувати взаємовпливи між рядами значень показників природного середовища.

Систематизовано усі розроблені методи розкриття видів невизначеності та запропоновано алгоритм системного аналізу стану атмосферного повітря з урахуванням аномалій. Алгоритм складається з 15 кроків та дозволяє комплексно проаналізувати систему управління станом природного середовища. Розкриття видів невизначеності, які були описані в підрозділі 1.6, наведено в додатку Д.

РОЗДІЛ 4

ОПТИМІЗАЦІЯ СИСТЕМИ МОНІТОРИНГУ ЗОНИ «ВІННИЦЬКА»

4.1. Системний аналіз стану атмосферного повітря у зоні «Вінницька»

З використанням елементів алгоритму, вказаного у підрозділі 3.4, у 2021 році на замовлення Управління розвитку територій та інфраструктури Вінницької обласної державної адміністрації та ТОВ «ІННОВІНПРОМ» здійснено оптимізацію системи моніторингу зони «Вінницька» (Вінницької області). Під час виконання цієї роботи використовувалися лише деякі елементи алгоритму, оскільки в роботі були чітко визначені вимоги і не було можливості застосувати усі попередні методи. Самі ж ці додаткові методи були розкриті окремо після виконання цієї роботи в рамках різного роду аналізу та прогнозування стану атмосферного повітря з використанням реальних даних моніторингу, приклади яких наведено у відповідних розділах цієї роботи.

Згідно алгоритму системного аналізу стану атмосферного повітря, одними з перших кроків є інформаційні запити на отримання комплексної інформації щодо стану системи, яка аналізується, а також переліку важливих для системного аналізу даних: відомі джерела забруднення атмосферного повітря, дозволи на викиди забруднювальних речовин, поточний стан мережі моніторингу стану атмосферного повітря, показники за якими відбувається моніторинг, тощо. Для розкрити ряду інформаційних видів невизначеності (U_{I1} , U_{I2} , U_{I3} , U_{I4} , U_{I5}), отримані дані необхідно проаналізувати.

За наявною відритою інформацією було охарактеризовано досліджувану зону, що включало базовий аналіз географічних, метеорологічних та інших показників. Результати цього аналізу описано у додатку Е. Також було проаналізовано основні джерела забруднення атмосферного повітря та обсяги цих забруднень у розрізі останніх декількох років – результати аналізу викладено в додатку Є.

Було проаналізовано поточний стан наявної мережі моніторингу якості атмосферного повітря. Станом на 2021 рік на території зони «Вінницька» існувала така мережа постів (табл. 4.1) та пунктів (табл. 4.2) спостережень за станом атмосферного повітря.

Таблиця 4.1 – Мережа постів спостережень за станом атмосферного повітря

Місце розташування посту або маршрути – точки відбору	Найменування юридичної особи	Перелік забруднювачів	Режим спостережень
Вінницький обласний лабораторний центр Міністерства охорони здоров'я України. Дослідження стану атмосферного повітря здійснюються на межі санітарно - захисної зони, у селітебній зоні, впродовж автомагістралей	Вінницький обласний лабораторний центр Міністерства охорони здоров'я України.	Азоту діоксид	1 раз в квартал
		Сірки діоксид	1 раз в квартал
		Аміак	1 раз в квартал
		Пил	1 раз в квартал
		Оксид вуглецю	1 раз в квартал
		Формальдегід	1 раз в квартал

Таблиця 4.2 – Мережа пунктів спостережень за станом атмосферного повітря

Місце розташування пункту	Найменування юридичної особи,	Перелік показників та складових опадів	Режим спостережень
Вінницький обласний центр з гідрометеорології			
Метеостанція Хмільник. Адреса: вул. Північна, 42, м. Хмільник, Хмільницький район Вінницької області широта 49°34', довгота 27°56'	Державна установа "Вінницький обласний центр з гідрометеорології ДСНС України". Адреса: вул. С. Руданського / І. Богуна буд. 14/127 м. Вінниця	Кислотність опадів за активністю іонів водню.	Відбір проб кожних атмосферних опадів для визначення вмісту їх складових
Метеостанція Жмеринка. Адреса: вул. Доватора, 125, м. Жмеринка, Жмеринський район Вінницької області широта 49°01', довгота 28°08'		Кислотність опадів за активністю іонів водню.	
Авіаційна метеорологічна станція цивільна Вінниця (АМСЦ Вінниця). Адреса: с. Гавришівка, 125, Вінницький район Вінницької області широта 49°22', довгота 28°60'		Кислотність опадів за активністю іонів водню.	Відбір проби снігового покриву: 1 раз на рік

Державний моніторинг якості атмосферного повітря зони «Вінниця» проводиться районними лабораторіями Вінницького обласного лабораторного центру МОЗ України.

Пункти Вінницького обласного лабораторного центру МОЗ України враховують сумарне забруднення від усіх джерел викидів, що розташовані в певному районі. Основними забруднювальними речовинами, для яких характерні перевищення допустимих гігієнічних нормативів якості атмосферного повітря, є такі:

- пил;
- діоксид сірки;
- діоксид азоту;
- оксид вуглецю;
- формальдегід;
- аміак.

Оцінювання стану атмосферного повітря у 2016-2021 роках на території зони «Вінницька» здійснювалось шляхом порівняння середньомісячних концентрацій з порогоми оцінювання відповідно до Постанови КМУ від 14.08.2019 р. № 827 [10] та з середньодобовими гранично-допустимими концентраціями (далі – ГДК) по 7 пріоритетних забруднювальних речовинах, які найбільше впливають на забруднення атмосферного повітря. За відсутності порогів оцінювання для певних речовин оцінювання проводилося лише шляхом порівняння середньомісячних концентрацій з середньодобовим ГДК. Для пилу (недиференційованого за складом) оцінювання проводилося теж лише шляхом порівняння середньомісячних концентрацій з середньодобовою ГДК, та порівняння максимально разових концентрацій з максимально разовою ГДК, оскільки верхні пороги оцінювання, визначені у Постанові КМУ від 14.08.2019 р. № 827 [10] існують тільки окремо для ТЧ₁₀ та ТЧ_{2.5}. Перелік пріоритетних забруднювальних речовин наведено у таблиці з вказанням їх ГДК (таблиця 4.3).

Таблиця 4.3. Перелік пріоритетних забруднювальних речовин з їх ГДК

Забруднювальна речовина	Середньодобова (мг/м ³)	Максимально разова (мг/м ³)
Пил	0,15	0,5
Діоксид сірки	0,05	0,5
Діоксид азоту	0,04	0,2
Оксид вуглецю	3	5
Оксид азоту	0,06	0,4
Формальдегід	0,003	0,035
Аміак	0,004	0,2

В результаті порівняння результатів спостережень за концентрацією діоксиду сірки із верхнім та нижнім порогоми оцінювання (рис. 4.1) виявлено постійне перевищення нижнього порогу оцінювання та одиничні випадки перевищень верхнього порогу оцінювання.

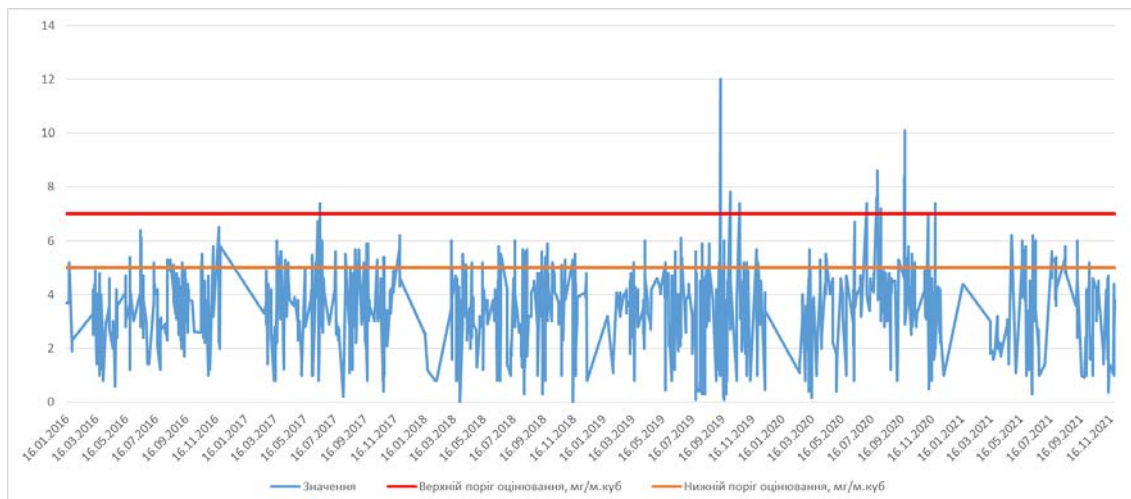


Рисунок 4.1 – Порівняння результатів спостережень діоксиду сірки з порогами оцінювання

В результаті порівняння результатів спостережень за концентрацією діоксиду азоту із ГДК (рис. 4.2) виявлено постійне перевищення середньодобового ГДК, та періодичні перевищення максимально разової ГДК, найбільшими з яких за останній час є такі: у вересні 2019 року та в липні 2020 року у смт Тростянець та у вересні 2020 року у смт Турбів.

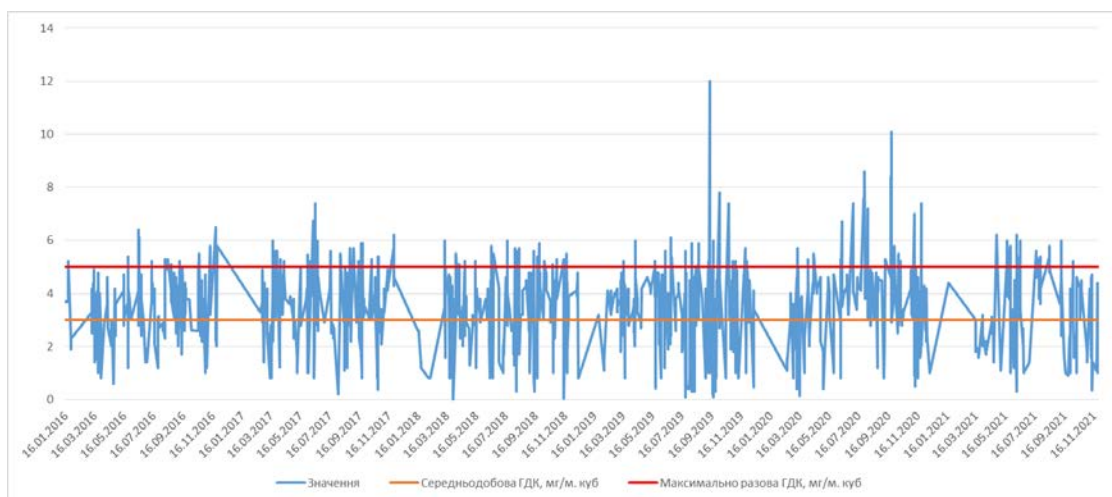


Рисунок 4.2 – Порівняння результатів спостережень оксиду вуглецю із ГДК

В результаті порівняння результатів спостережень за концентрацією діоксиду сірки із верхнім та нижнім порогами оцінювання (рис. 4.3) виявлено постійне перевищення як верхнього, так і нижнього порогів оцінювання.

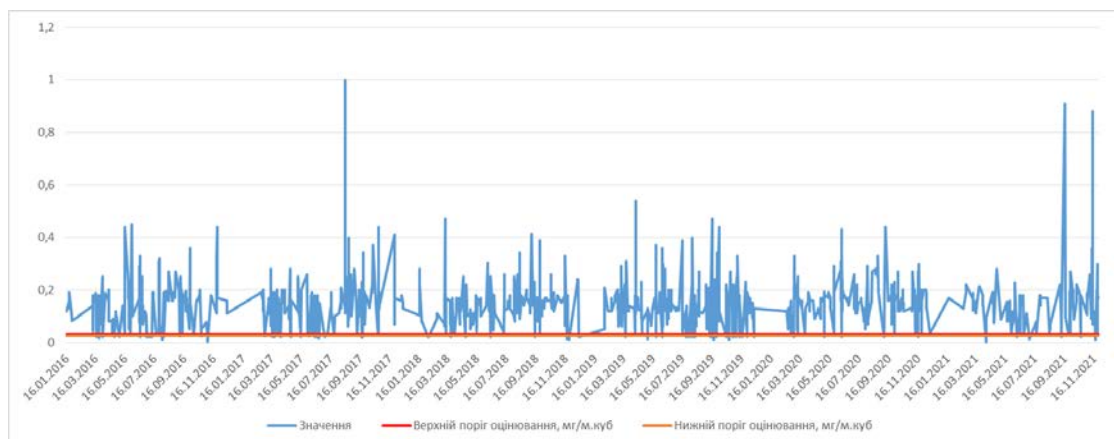


Рисунок 4.3 – Порівняння результатів спостережень діоксиду азоту з порогоми оцінювання

В результаті порівняння результатів спостережень за концентрацією діоксиду азоту із ГДК (рис. 4.4) виявлено постійне перевищення середньодобового ГДК, та періодичні перевищення максимально разової ГДК, найбільшими з яких за останній час є такі: у вересні 2021 року в с. Іванів, в листопаді 2021 року у м. Калинівка.

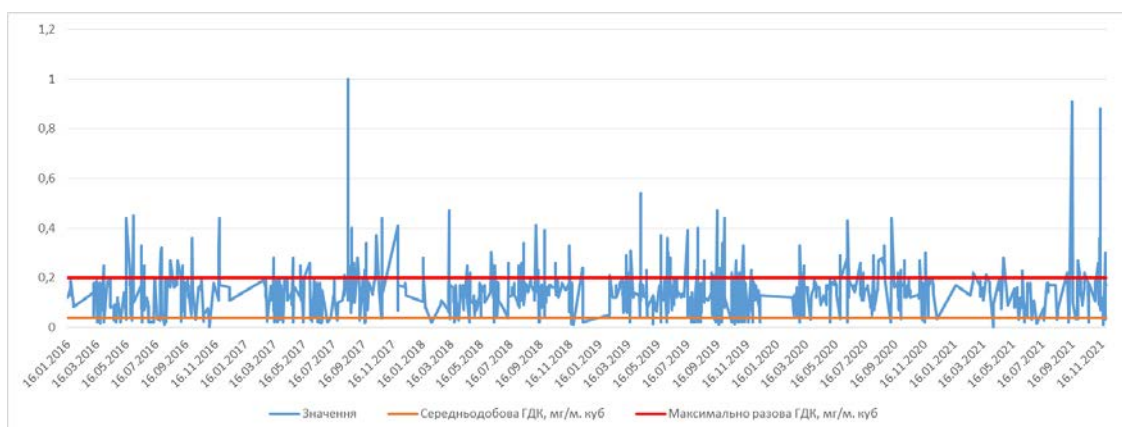


Рисунок 4.4 – Порівняння результатів спостережень діоксиду азоту із ГДК

В результаті порівняння результатів спостережень за концентрацією діоксиду сірки із верхнім та нижнім порогоми оцінювання (рис. 4.5) виявлено дуже часті перевищення як верхнього, так і нижнього порогів оцінювання.

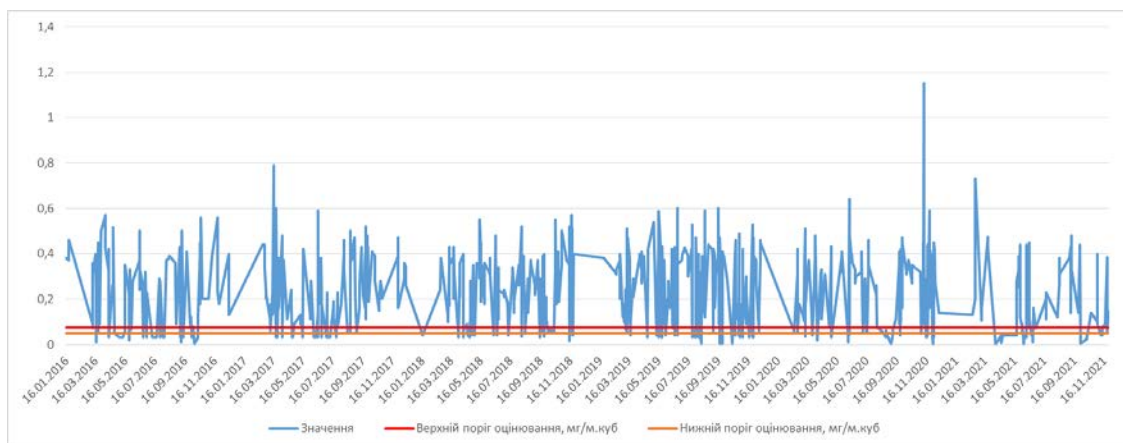


Рисунок 4.5 – Порівняння результатів спостережень діоксиду сірки з порогоми оцінювання

В результаті порівняння результатів спостережень за концентрацією діоксиду сірки із ГДК (рис. 4.6) виявлено дуже часті перевищення середньодобового ГДК та періодичні перевищення максимальної разової ГДК, найбільшими з яких за останній час є такі: в червні 2020 року в м. Жмеринка, в листопаді 2020 року в м. Бершадь та в лютому 2021 року в м. Козятин.

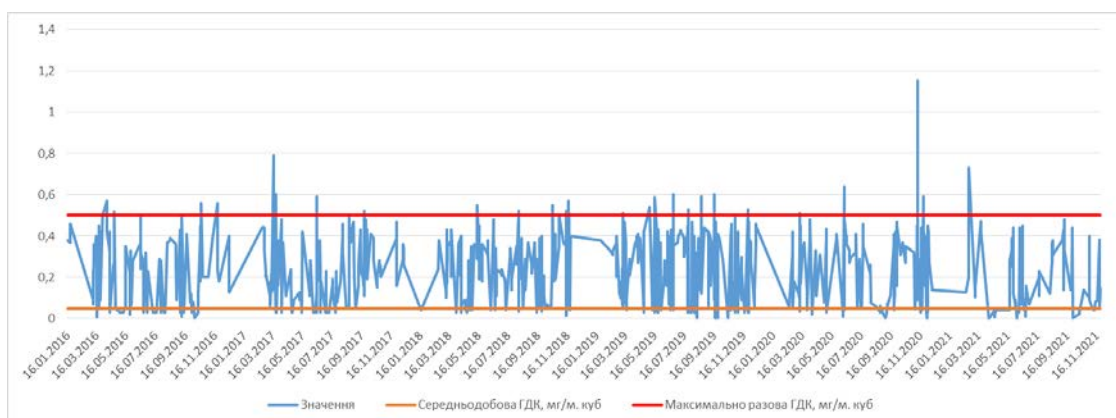


Рисунок 4.6 – Порівняння результатів спостережень діоксиду сірки із ГДК

В результаті порівняння результатів спостережень за концентрацією аміаку із ГДК (рис. 4.7) виявлено дуже часті перевищення середньодобового ГДК та одиночні перевищення максимальної разової ГДК: у вересні 2016 року в м. Немирів та у березні 2018 року в м. Тульчин.

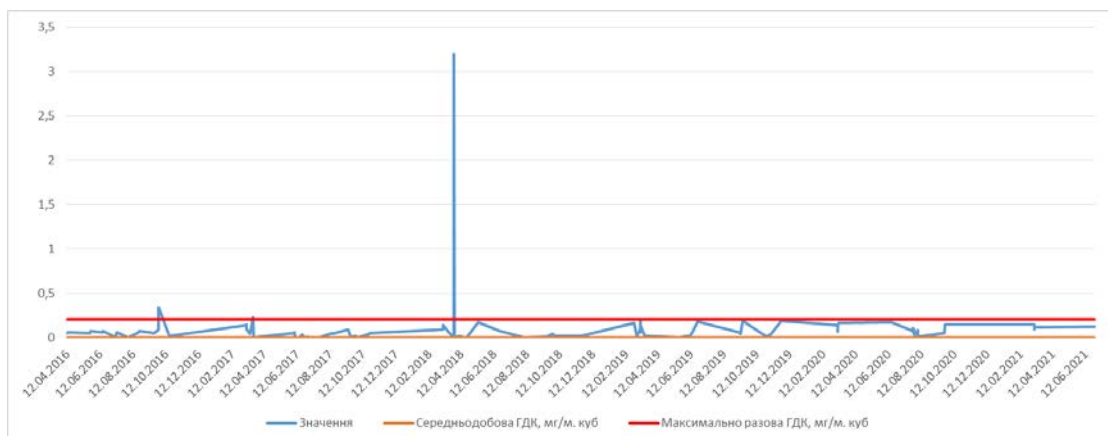


Рисунок 4.7 – Порівняння результатів спостережень аміаку із ГДК

В результаті порівняння результатів спостережень за концентрацією пилу із ГДК (рис. 4.8) виявлено постійне перевищення середньодобового ГДК, та періодичні перевищення максимально разової ГДК. Більшість з цих перевищень була виявлена до середини 2017 року. За останні три роки суттєві перевищення виявлені у серпні та листопаді 2019 року в м. Жмеринка.

Оскільки показники концентрації пилу (PM2.5, PM10) є одними із ключових показників, які використовуються для визначення мінімальної кількості пунктів спостережень у зоні, а аналіз недиференційованого пилу виявив суттєві перевищення обох видів ГДК, то були проаналізовані й дані громадського моніторингу атмосферного повітря зони. Щохвилинні дані були імпортовані з порталу SaveEcoBot <https://www.saveecobot.com/> за 2019-2021 рр. Крім того, були отримані щохвилинні дані з мережі моніторингу «EcoCity» [18] (у межах Меморандуму від громадської організації «Фрі-Ардуіно» та Вінницьким національним технічним університетом, який є одним з розробників даної Програми). Дані були усереднені у щодобові. Аналіз показав, що перевищення верхнього порогу трапляються досить часто, але остаточний висновок варто буде зробити після встановлення та використання автоматизованих постів державної системи моніторингу по цих показниках. А поки варто просто взяти це до уваги.

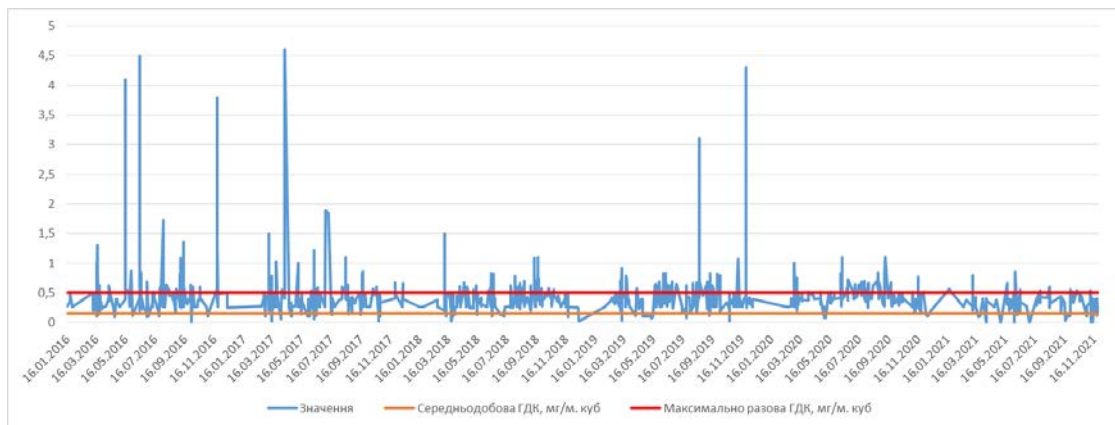


Рисунок 4.8 – Порівняння результатів спостережень пилу із ГДК

В результаті порівняння результатів спостережень формальдегіду із ГДК (рис. 4.9) виявлено дуже часті перевищення середньодобового ГДК та періодичні перевищення максимально разової ГДК, зокрема такі: у листопаді 2018 року та у травні 2019 року у смт Муровані Курилівці, у червні 2018 року у смт Копайгород, у березні 2020 року в с. Війтівці Хмельницького району.

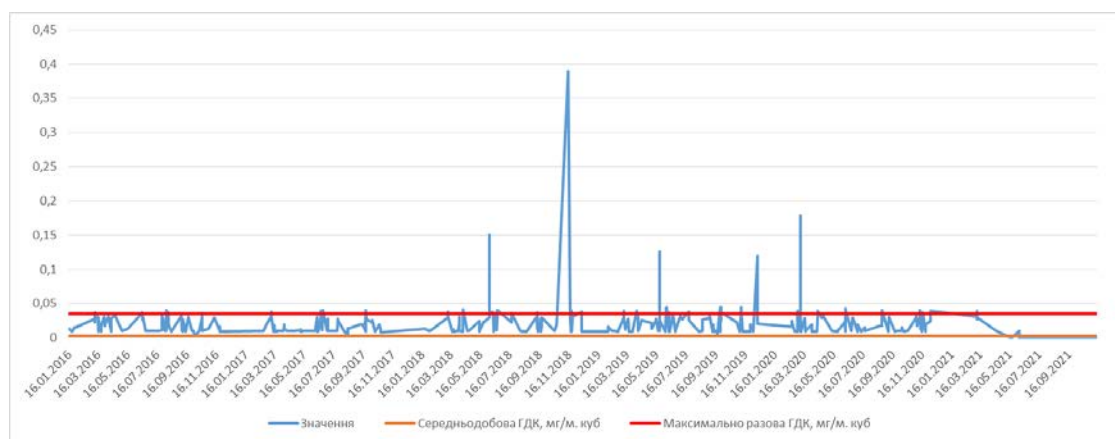


Рисунок 4.9 – Порівняння результатів спостережень формальдегіду із ГДК

Після встановлення постів автоматизованого спостереження та отримання фактичних даних щодо забруднення атмосферного повітря будуть внесені зміни до Програми у частині зміни режиму оцінювання. Згідно п.9 Порядку режим оцінювання, крім режиму фіксованих вимірювань, може бути переглянутий раніше за рішенням органів управління якістю атмосферного повітря за таких умов:

- 1) режим оцінювання із застосуванням методів моделювання та об'єктивного оцінювання замінюється на режим комбінованого оцінювання;
- 2) режим комбінованого оцінювання замінюється на режим фіксованих вимірювань.

Згідно п.21 Порядку, якщо суб'єктами моніторингу атмосферного повітря заплановано додаткові заходи щодо встановлення пунктів спостережень та/або вдосконалення наявних мереж спостереження за якістю атмосферного повітря, створення та/або вдосконалення лабораторій спостереження за станом атмосферного повітря після затвердження програми, органи управління якістю атмосферного повітря вносять зміни до програми за поданням таких суб'єктів.

На рисунку 4.10 наведено карту просторового розподілу концентрацій діоксиду сірки, побудовану шляхом просторової інтерполяції усереднених результатів спостережень за станом атмосферного повітря за 2020 рік методом середньозважених відстаней.

Найбільший рівень забруднення атмосфери діоксидом сірки спостерігається на території м. Іллінці, смт Оратів, смт Турбів, м. Жмеринка, м. Бар, м. Шаргород, смт Теплик, смт Чечельник та м. Ладижин.

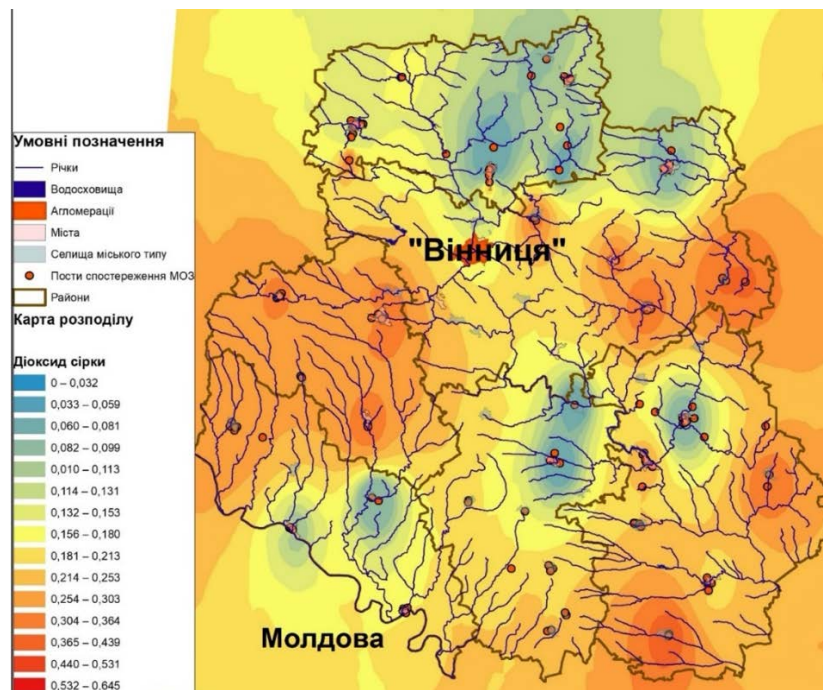


Рисунок 4.10 – Карта просторового розподілу концентрацій діоксиду сірки

На рисунку 4.11 приведено карту просторового розподілу концентрацій оксиду вуглецю, побудовану шляхом просторової інтерполяції усереднених результатів спостережень за станом атмосферного повітря за 2020 рік методом середньозважених відстаней.

Найбільший рівень забруднення атмосфери оксидом вуглецю спостерігається на території м. Жмеринка, м. Шаргород, смт Тростянець та неподалік від м. Гайсин.

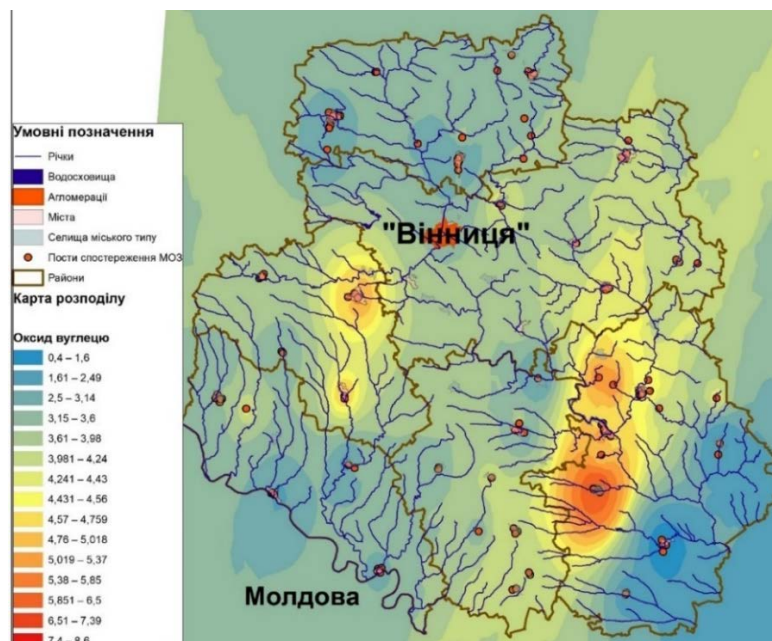


Рисунок 4.11 – Карта просторового розподілу концентрацій оксиду вуглецю

На рисунку 4.12 наведено карту просторового розподілу концентрацій пилу, побудовану шляхом просторової інтерполяції усереднених результатів спостережень за станом атмосферного повітря за 2020 рік методом середньозважених відстаней.

Найбільший рівень забруднення атмосфери пилом спостерігається на території м. Калинівка, м. Гайсин, с. Краснопілка, м. Могилів-Подільський.

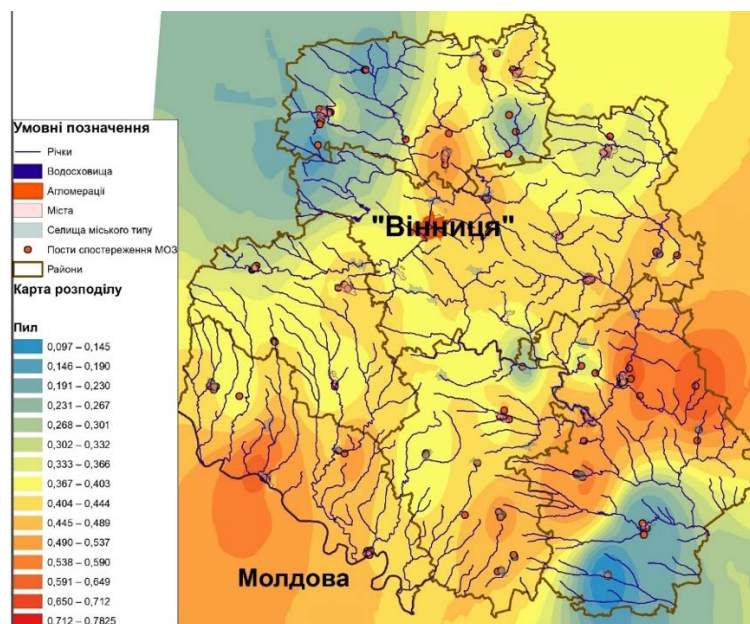


Рисунок 4.12 – Карта просторового розподілу пилу

На рисунку 4.13 наведено карту просторового розподілу концентрацій діоксиду азоту, побудовану шляхом просторової інтерполяції усереднених результатів спостережень за станом атмосферного повітря за 2020 рік методом середньозважених відстаней.

Найбільший рівень забруднення атмосфери діоксидом азоту спостерігається на території м. Калинівки. Також підвищений рівень діоксиду азоту спостерігається у м. Гайсин, м. Могилів-Подільський та у смт Чернівці, тобто на автострадах міжнародного значення та дорогах місцевого значення, які їх з'єднують (Т 0230).

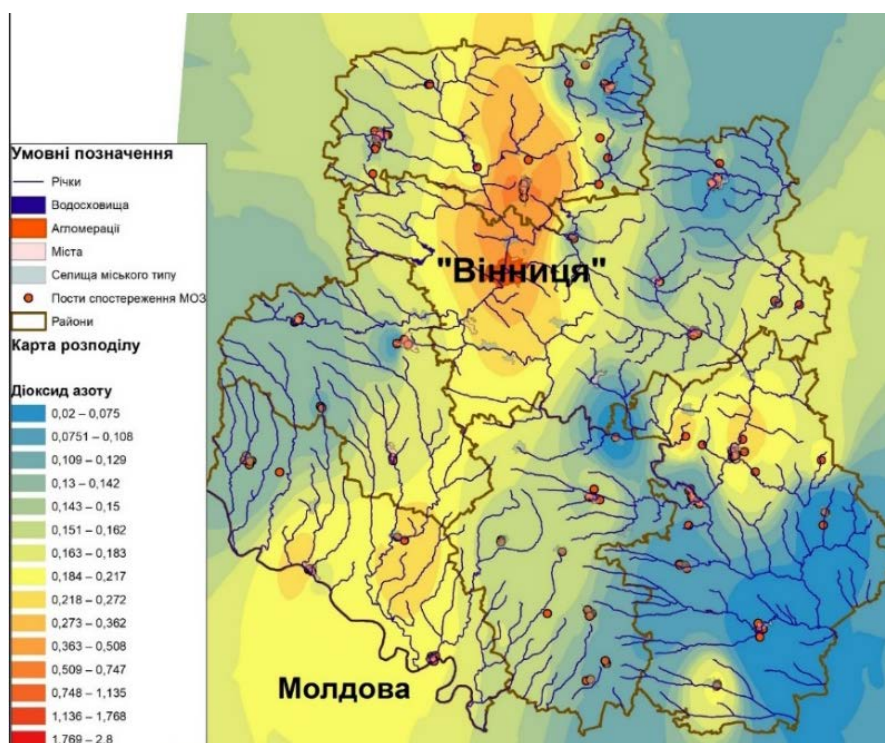


Рисунок 4.13 – Карта просторового розподілу концентрацій діоксиду азоту

За даними ДУ «Вінницький лабораторний центр МОЗ України» здійснено аналіз розподілу концентрацій забруднювальних речовин уздовж доріг зони «Вінницька». Результат аналізу подано у вигляді тематичної карти, яка підтверджує підвищені концентрації на пунктах, які розташовані поряд з основними дорогами міжнародного, національного, регіонального та місцевого значення, що проходять по території зони «Вінницька» (рис. 4.14).



Рисунок 4.14 – Карта розподілу концентрацій забруднювальних речовин уздовж доріг

В подальшому карти просторового розподілу концентрацій забруднювальних речовин та статистику розподілу концентрацій забруднювальних речовин уздовж доріг будуть готуватися на основі моніторингових даних постів автоматизованих спостережень, що будуть розташовані на території зони «Вінницька».

На рисунку 4.15 наведено графік динаміки сумарних викидів забруднювальних речовин від стаціонарних джерел на території зони «Вінницька».



Рисунок 4.15 – Викиди забруднювальних речовин від стаціонарних джерел (тис. т.) зони «Вінницька»

Результати аналізу узагальнені у додатку Ж.

4.2. Оптимізація мережі спостережень та оцінювання методом системного аналізу

На території зони «Вінницька» відсутні автоматизовані пости спостережень за якістю атмосферного повітря та опадів.

Відповідно до вимог наказу МВС України від 21.04.2021 р. № 300 «Порядок розміщення пунктів спостережень за забрудненням атмосферного повітря в зонах та агломераціях» [12] зона «Вінницька» з населенням 1 147,477 тис. жителів повинна мати мінімум 4 пункти спостережень з фіксованими вимірюваннями. Аналіз показників по концентрації пилу (і порівняння недиференційованого пилу з ГДК за даними державного моніторингу, і порівняння з верхнім порогом даних окремо по PM2.5 та PM10 за даними громадського моніторингу) показав наявність частих і значних перевищень, що означає необхідність спостережень за мінімально 6, а не 4, пунктами спостережень. Детальніше про це в додатку 3

Під час вибору місць розташування пунктів спостережень системи моніторингу атмосферного повітря використовувались методи системного аналізу, зокрема, метод аналізу ієрархій, з урахуванням таких критеріїв:

Φ₁ – близькість до найбільш впливових джерел забруднення атмосферного повітря;

Φ₂ – частота та рівень перевищень усереднених та максимальних разових показників стану атмосферного повітря та атмосферних опадів (топографічні та метеорологічні дані, що можуть впливати на особливості розсіювання забруднювальних речовин);

Φ₃ – близькість до стейкхолдерів (статистичні відомості про щільність населення, інтенсивність найближчих транспортних потоків);

Φ₄ – рівень репрезентативності запроваджених керівних заходів (встановлені очисні системи, введений контроль за викидами пересувних транспортних заходів тощо) (група R-критеріїв – «Responses»);

Φ₅ – зручність інфраструктури, умов для обслуговування пунктів спостережень, цінність для рівномірного охоплення регіону, що спростить побудову карт забруднення в майбутньому та ін.

З метою впровадження державної системи спостережень за станом атмосферного повітря на території зони «Вінницька», Вінницькою обласною державною адміністрацією було запропоновано першочергово до 2022 року придбання та встановлення 4 автоматизованих еталонних постів спостереження за якістю атмосферного повітря в м. Калинівка, м. Іллінці, м. Гайсин та м. Ладижин, а також 2 автоматизованих індикативних постів спостереження за якістю атмосферного повітря в м. Хмільнику та в с. Немія Могилів-Подільського району.

В подальшому протягом 2023-2025 років запропоновано придбати

1 пересувну лабораторію контролю якості атмосферного повітря, яка могла б охопити населені пункти, де були виявлені найбільші концентрації забруднювальних речовин, у т.ч. на мережі громадського моніторингу атмосферного повітря, але які не увійшли до переліку рекомендованих місць для встановлення стаціонарних пунктів: с. Іванів, м. Шаргород, м. Могилів-Подільський, смт Оратів, смт Тростянець, смт Турбів, м. Жмеринка, м. Козятин, м. Бершадь, м. Тульчин, м. Немирів, смт Муровані Курилівці, смт Копайгород, с. Війтівці Хмельницького району, м. Бар, смт Теплик, смт Чечельник, с. Краснопілка, смт Чернівці.

Проектування місць розташування запланованих Програмою автоматизованих постів спостережень, з урахуванням пропозицій Комісії з питань здійснення державного моніторингу в галузі охорони атмосферного повітря та управління якістю атмосферного повітря, здійснювалось у відповідності до вимог Порядку здійснення державного моніторингу в галузі охорони атмосферного повітря, затвердженого Постановою та наказу Міністерства внутрішніх справ України від 21.04.2021 № 300 «Про затвердження Порядку розміщення пунктів спостережень за забрудненням атмосферного повітря в зонах та агломераціях». Карта (рис. 4.16) запланованих місць розміщення пунктів спостереження за станом атмосферного повітря зони «Вінницька» додається.

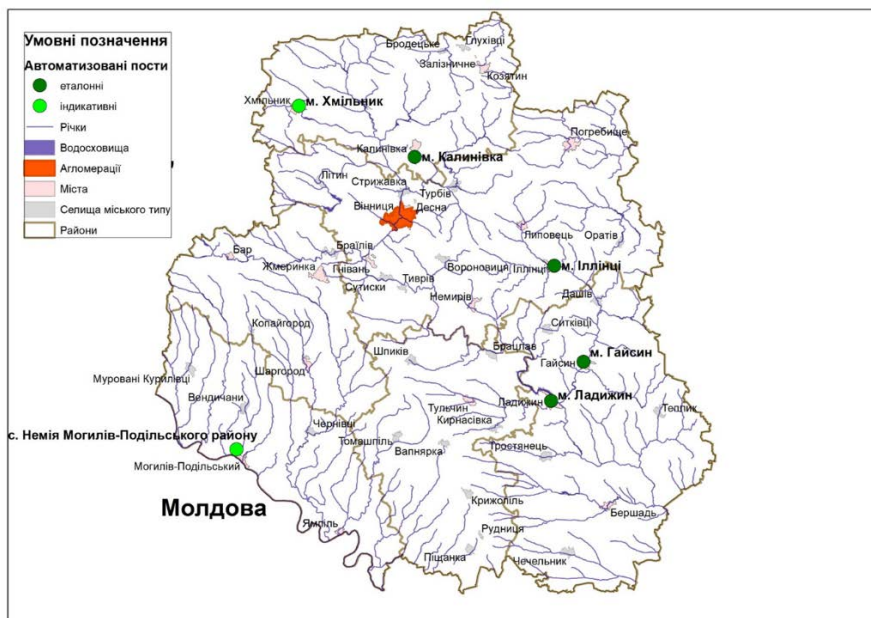


Рисунок 4.16 – Карта місць розміщення пунктів спостереження за станом атмосферного повітря зони «Вінницька»

В таблиці 4.4 приведено розміщення та кількість пунктів спостереження (по постах).

Таблиця 4.4 – Розміщення та кількість пунктів спостереження (по постах)

№	Місце розташування пункту спостережень (адреса/координати) або маршрут	Тип пункту спостережень	Перелік забруднювальних речовин
1.	Автоматизований пост м. Калинівка, біля перехрестя вул. Механізаторів та автомагістралі М21 49.423065, 28.524014	міський транспортний	Діоксид сірки Діоксид азоту та оксиди азоту Оксид вуглецю Тверді частки (ТЧ ₁₀) Тверді частки (ТЧ _{2,5}) Бензол Бенз(а)пірен Озон
2.	Автоматизований пост м. Іллінці, біля перехрестя вул. Незалежності та вул. Коцюбинського 49.107293, 29.217228	міський промисловий	Діоксид сірки Діоксид азоту та оксиди азоту Оксид вуглецю Тверді частки (ТЧ ₁₀), Тверді частки (ТЧ _{2,5}), Озон
3.	Автоматизований пост м. Гайсин, біля перехрестя вул. Жовтневої та вул. Плеханова 48.811574, 29.377752	міський промисловий	Діоксид сірки Діоксид азоту та оксиди азоту Тверді частки (ТЧ ₁₀) Тверді частки (ТЧ _{2,5}) Озон Бензол Оксид вуглецю Бенз(а)пірен <i>Показники опадів:</i> Кількість опадів Іони амонію Гідрокарбонат-іони Іони калію Іони кальцію Загальна кислотність Іони магнію Іони натрію Нітрат-іони Сульфат-іони Хлорид-іони рН

Продовження таблиці 4.4

№	Місце розташування пункту спостережень (адреса/координати) або маршрут	Тип пункту спостережень	Перелік забруднювальних речовин
4.	Автоматизований пост м. Гайсин, біля перехрестя вул. Жовтневої та вул. Плеханова 48.811574, 29.377752	міський промисловий	Діоксид сірки Діоксид азоту та оксиди азоту Тверді частки (ТЧ10) Тверді частки (ТЧ2,5) Озон Бензол Оксид вуглецю Бенз(а)пірен <i>Показники та складові атмосферних опадів:</i> Кількість опадів Іони амонію Гідрокарбонат-іони Іони калію Іони кальцію Загальна кислотність Іони магнію Іони натрію Нітрат-іони Сульфат-іони Хлорид-іони рН
5.	Автоматизований пост м. Ладижин, перехрестя вул. Петра Кравчика та вул. Будівельника, 84 48.684991, 29.234384	міський промисловий	Діоксид сірки Діоксид азоту та оксиди азоту Оксид вуглецю Тверді частки (ТЧ10), Тверді частки (ТЧ2,5), Озон Арсен Свинець Кадмій Нікель Ртуть
6.	Пересувна лабораторія контролю якості атмосферного повітря	міський/сільський фоновий	Діоксид сірки Діоксид азоту та оксиди азоту Оксид вуглецю Тверді частки (ТЧ ₁₀) Тверді частки (ТЧ _{2,5}), Озон

Продовження таблиці 4.4

№	Місце розташування пункту спостережень (адреса/координати) або маршрут	Тип пункту спостережень	Перелік забруднювальних речовин
7.	Автоматизований пост с. Немія Могилів-Подільського району 48.475623, 27.769802	сільський фоновий	Діоксид сірки Діоксид азоту та оксиди азоту Оксид вуглецю Тверді частки (ТЧ ₁₀) Тверді частки (ТЧ _{2,5}), Озон
8.	Автоматизований пост м. Хмільник біля перехрестя вул. Пушкіна та вул. Небесної Сотні 49.559058, 27.954836	міський фоновий	Діоксид сірки Діоксид азоту та оксиди азоту Оксид вуглецю Тверді частки (ТЧ ₁₀) Тверді частки (ТЧ _{2,5}), Озон

При проєктуванні мережі спостережень за станом атмосферного повітря враховано встановлений режим оцінювання для забруднювальних речовин у зоні та положення «Порядку розміщення пунктів спостережень за забрудненням атмосферного повітря в зонах та агломераціях», затвердженого наказом Міністерства внутрішніх справ України від 21 квітня 2021 року № 300 (zareєстровано в Міністерстві юстиції України 13 травня 2021 р. за N 635/36257).

У разі уточнення режимів оцінювання для кожної забруднювальної речовини розміщення та кількість пунктів автоматизованого спостереження за станом атмосферного повітря може бути скоригована, шляхом внесення відповідних змін до Програми.

В таблиці 4.5 приведено розміщення та кількість пунктів спостереження (по забруднювальних речовинах).

Таблиця 4.5 – Розміщення та кількість пунктів спостереження (по забруднювальних речовинах).

№	Забруднювальна речовина	Тип станції	Вид вимірювань	Мета досліджень	Тип території та місце розташування пункту
1.	Діоксид сірки	транспортна	фіксовані автоматизовані	охорона здоров'я	Тип території - міська пост м. Калинівка, біля перехрестя вул. Механізаторів та автомагістралі М21 49.423065, 28.524014
2.	Діоксид азоту та оксиди азоту	транспортна	фіксовані автоматизовані	охорона здоров'я	
3.	Оксид вуглецю	транспортна	фіксовані автоматизовані	охорона здоров'я	
4.	Тверді частки (ТЧ ₁₀)	транспортна	фіксовані автоматизовані	охорона здоров'я	
5.	Тверді частки (ТЧ _{2,5})	транспортна	фіксовані автоматизовані	охорона здоров'я	
6.	Озон	транспортна	фіксовані автоматизовані	охорона здоров'я	
7.	Бензол	транспортна	фіксовані автоматизовані	охорона здоров'я	
8.	Бенз(а)пірен	транспортна	фіксовані автоматизовані	охорона здоров'я	
9.	Діоксид сірки	промислова	фіксовані автоматизовані	охорона здоров'я	Тип території - міська Пост м. Іллінці, біля перехрестя вул. Незалежності та вул. Коцюбинського 49.107293, 29.217228
10.	Діоксид азоту та оксиди азоту	промислова	фіксовані автоматизовані	охорона здоров'я	
11.	Оксид вуглецю	промислова	фіксовані автоматизовані	охорона здоров'я	
12.	Тверді частки (ТЧ ₁₀)	промислова	фіксовані автоматизовані	охорона здоров'я	
13.	Тверді частки (ТЧ _{2,5})	промислова	фіксовані автоматизовані	охорона здоров'я	
14.	Озон	промислова	фіксовані автоматизовані	охорона здоров'я	

Продовження таблиці 4.5

№	Забруднювальна речовина	Тип станції	Вид вимірювань	Мета досліджень	Тип території та місце розташування пункту
15.	Діоксид сірки	промислова	фіксовані автоматизовані	охорона здоров'я	Тип території - міська пост м. Гайсин, біля перехрестя вул. Жовтневої та вул. Плеханова 48.811574, 29.377752
16.	Діоксид азоту та оксиди азоту	промислова	фіксовані автоматизовані	охорона здоров'я	
17.	Оксид вуглецю	промислова	фіксовані автоматизовані	охорона здоров'я	
18.	Тверді частки (ТЧ ₁₀)	промислова	фіксовані автоматизовані	охорона здоров'я	
19.	Тверді частки (ТЧ _{2,5})	промислова	фіксовані автоматизовані	охорона здоров'я	
20.	Озон	промислова	фіксовані автоматизовані	охорона здоров'я	
21.	Бензол	промислова	фіксовані автоматизовані	охорона здоров'я	
22.	Бенз(а)пірен	промислова	фіксовані автоматизовані	охорона здоров'я	
23.	Діоксид сірки	промислова	фіксовані автоматизовані	охорона здоров'я	Тип території - міська пост м. Ладижин, перехрестя вул. Петра Кравчика та вул. Будівельника, 84 48.684991, 29.234384
24.	Діоксид азоту та оксиди азоту	промислова	фіксовані автоматизовані	охорона здоров'я	
25.	Оксид вуглецю	промислова	фіксовані автоматизовані	охорона здоров'я	
26.	Тверді частки (ТЧ ₁₀)	промислова	фіксовані автоматизовані	охорона здоров'я	
27.	Тверді частки (ТЧ _{2,5})	промислова	фіксовані автоматизовані	охорона здоров'я	
28.	Озон	промислова	фіксовані автоматизовані	охорона здоров'я	
29.	Арсен	промислова	фіксовані автоматизовані	охорона здоров'я	
30.	Свинець	промислова	фіксовані автоматизовані	охорона здоров'я	

Продовження таблиці 4.5

№	Забруднювальна речовина	Тип станції	Вид вимірювань	Мета досліджень	Тип території та місце розташування пункту
31.	Кадмій	промислова	фіксовані автоматизовані	охорона здоров'я	
32.	Нікель	промислова	фіксовані автоматизовані	охорона здоров'я	
33.	Ртуть	промислова	фіксовані автоматизовані	охорона здоров'я	
34.	Діоксид сірки	фонова	індикативні автоматизовані	охорона здоров'я	Пересувна лабораторія контролю якості атмосферного повітря
35.	Діоксид азоту та оксиди азоту	фонова	індикативні автоматизовані	охорона здоров'я	
36.	Оксид вуглецю	фонова	індикативні автоматизовані	охорона здоров'я	
37.	Тверді частки (ТЧ ₁₀)	фонова	індикативні автоматизовані	охорона здоров'я	
38.	Тверді частки (ТЧ _{2,5})	фонова	індикативні автоматизовані	охорона здоров'я	
39.	Озон	фонова	індикативні автоматизовані	охорона здоров'я	Тип території - сільська пост с. Немія Могилів-Подільського району 48.475623, 27.769802
40.	Діоксид сірки	фонова	індикативні автоматизовані	охорона здоров'я	
41.	Діоксид азоту та оксиди азоту	фонова	індикативні автоматизовані	охорона здоров'я	
42.	Оксид вуглецю	фонова	індикативні автоматизовані	охорона здоров'я	
43.	Тверді частки (ТЧ ₁₀)	фонова	індикативні автоматизовані	охорона здоров'я	
44.	Тверді частки (ТЧ _{2,5})	фонова	індикативні автоматизовані	охорона здоров'я	
45.	Озон	фонова	індикативні автоматизовані	охорона здоров'я	Тип території - міська пост м. Хмільник біля перехрестя вул. Пушкіна та вул. Небесної Сотні 49.559058, 27.954836
46.	Діоксид сірки	фонова	індикативні автоматизовані	охорона здоров'я	
47.	Діоксид азоту та оксиди азоту	фонова	індикативні автоматизовані	охорона здоров'я	
48.	Оксид вуглецю	фонова	індикативні автоматизовані	охорона здоров'я	
49.	Тверді частки (ТЧ ₁₀)	фонова	індикативні автоматизовані	охорона здоров'я	
50.	Тверді частки (ТЧ _{2,5})	фонова	індикативні автоматизовані	охорона здоров'я	
51.	Озон	фонова	індикативні автоматизовані	охорона здоров'я	

Застосування технології системного аналізу до оптимізації системи моніторингу довкілля Вінницької області, спільно з експертами підрозділів органів влади та місцевого самоврядування, дозволило розробити комплекс заходів щодо встановлення пунктів спостереження та/або вдосконалення наявних мереж спостереження за якістю атмосферного повітря, створення та/або вдосконалення лабораторій спостереження за станом атмосферного повітря, який увійшов у Програму моніторингу і був затверджений відповідними органами влади та місцевого самоврядування. Повний текст наведено у Додатку І.

4.3. Висновки до розділу 4

В даному розділі наведено огляд результатів апробації розроблених авторами методів та засобів системного аналізу стану атмосферного повітря. Описано впровадження результатів для оптимізації системи моніторингу повітря в межах зони «Вінницька» (Вінницької області).

ВИСНОВКИ

У результаті виконання даної роботи розв'язано актуальну науково-прикладну задачу системного аналізу стану атмосферного повітря регіону, з урахуванням впливу аномалій, що дозволяє комплексно підходити до проблем, що мають місце при функціонуванні систем управління та моніторингу станом атмосферного повітря та дозволяє покращувати їх ефективність. У ході розв'язання поставлених задач отримано такі наукові та науково-практичні результати:

1. Проаналізовано сучасний стан атмосферного повітря в Україні та які існують державні механізми контролю за цим станом. Проаналізовано основні показники якості атмосферного повітря. Здійснено огляд європейського індексу якості повітря та порівняно з українським аналогом. Описано мережі державного та громадського моніторингу, яку функціонують на території України.

2. Описано загальну методологію системного аналізу та згідно неї охарактеризовано систему управління станом атмосферного повітря, проведено декомпозицію системи за різними варіантами декомпозиції та виділено види невизначеності, що мають місце в таких системах.

3. Проведено порівняльний аналіз методів та інтелектуальних моделей аналізу та прогнозування часових рядів стану атмосферного повітря. Проведено порівняльний аналіз відомих методів пошуку аномалій в часових рядах та засоби автоматизації цих методів з використанням мови програмування Python.

4. Систематизовано методи розкриття видів невизначеності та визначено які методи та підходи потребують розробки або удосконалення.

5. Запропоновано метод пошуку можливих джерел підвищеного антропогенного навантаження на стан атмосферного повітря, що дозволяє ідентифікувати місце та час, де розпочалося стале зростання забруднення.

6. Запропоновано підхід для формування інтегральних критеріїв вибору місць розташування станцій моніторингу стану атмосферного повітря, що враховує аномальні та пікові значення.

7. Розроблено та удосконалено методи аналізу та прогнозування стану атмосферного повітря з урахуванням аномалій. Ці методи дозволяють ідентифікувати різні параметри сезонності та аномалій часового ряду, що дозволить побудувати ефективну модель прогнозування.

8. Розроблено алгоритм системного аналізу стану атмосферного повітря з урахуванням аномалій, що базується на використанні класичних та авторських методів розкриття видів невизначеності, що мають місце в системі управління станом атмосферного повітря.

9. З використанням розроблених методів та алгоритму системного аналізу стану атмосферного повітря здійснено оптимізацію системи моніторингу зони «Вінницька» (території Вінницької області, яка не включає в себе місто Вінницю).

ПЕРЕЛІК ЛІТЕРАТУРИ

1. В. Є. Копняк, В. Б. Мокін, С. О. Жуков, І. В. Варчук, і Т. В. Скринник, «Метод бустингу гетероскедастичних моделей для прогнозування концентрацій пилу Сахари в атмосферному повітрі України», Наукові праці ВНТУ, вип. 2, Лип. 2024. <https://doi.org/10.31649/2307-5376-2024-2-28-38>.
2. Методологія та організація наукових досліджень: підручник – вид. 3-є, змін. та доп. [Електронний ресурс] / Б. І. Мокін, О. Б. Мокін, В. Б. Мокін. – Вінниця: ВНТУ, 2023. – (PDF, 230 с.)
3. Моделювання та оптимізація систем : підручник / Дубовой В. М., Кветний Р. Н., Михальов О. І., Усов А. В. Вінниця : ПП «ТД «Едельвейс», 2017. 804 с.
4. Згуровський М. З., Панкратова Н. Д. Основи системного аналізу. К. Видавнича група ВНУ, 2007. 546 с.
5. Наука про дані: машинне навчання та інтелектуальний аналіз даних: електронний навчальний посібник комбінованого (локального та мережевого) використання [Електронний ресурс] / В. Б. Мокін, М. В. Дратований – Вінниця : ВНТУ, 2024. – 258 с.
6. Typology of indicators and the DPSIR framework. [Електронний ресурс] – Режим доступу: <https://www.eea.europa.eu/publications/TEC25>.
7. Про затвердження Програми державного моніторингу в галузі охорони атмосферного повітря на 2023–2027 роки для зони «Вінницька». [Електронний ресурс] – Режим доступу: <https://mepr.gov.ua/wp-content/uploads/2023/01/Natsdopovid-2021-n.pdf>.
8. What is particulate matter and what are its effects on human health?. [Електронний ресурс] – Режим доступу: <https://www.eea.europa.eu/en/about/contact-us/faqs/what-is-particulate-matter-and-what-are-its-effects-on-human-health>.
9. Impacts of air pollution on ecosystems in Europe. [Електронний ресурс] – Режим доступу: <https://www.eea.europa.eu/en/analysis/publications/impacts-of-air-pollution-on-ecosystems-in-europe>.
10. Постанова Кабінету Міністрів України від 14.09.2019 № 827 «Деякі питання здійснення державного моніторингу у галузі охорони атмосферного повітря», якою затверджено Порядок здійснення державного моніторингу в галузі охорони атмосферного повітря [<https://zakon.rada.gov.ua/laws/show/827-2019-%D0%BF#Text>].
11. Наказ Міністерства захисту довкілля та природних ресурсів України від 25.02.2021 № 147 «Про затвердження форми Програми державного моніторингу у галузі охорони атмосферного повітря», зареєстрований у Міністерстві юстиції України 21.04.2021 за № 543/36165 [<https://zakon.rada.gov.ua/laws/show/z0543-21#n14>].

12. Наказ Міністерства внутрішніх справ України від 21.04.2021 № 300 «Про затвердження Порядку розміщення пунктів спостережень за забрудненням атмосферного повітря в зонах та агломераціях», зареєстрований у Міністерстві юстиції України 13.05.2021 за № 635/36257 [<https://zakon.rada.gov.ua/laws/show/z0635-21#Text>].

13. Закон України про навколишнє середовище [<https://zakon.rada.gov.ua/laws/show/1264-12#Text>].

14. Національна доповідь про стан навколишнього природного середовища в Україні у 2021 році. [Електронний ресурс] – Режим доступу: <https://mepr.gov.ua/wp-content/uploads/2023/01/Natsdopovid-2021-n.pdf>.

15. Екологічний паспорт Вінницької області за 2024 рік. [Електронний ресурс] – Режим доступу: https://www.vin.gov.ua/images/UPRTER/2024/ogoloshennya/Ekologichnu_y%20pasport%202024_%20.pdf.

16. Угода про асоціацію з ЄС [https://zakon.rada.gov.ua/laws/show/984_011#Text].

17. «Чисте повітря для України». [Електронний ресурс] – Режим доступу: <https://cleanair.org.ua/>.

18. «EcoCity». [Електронний ресурс] – Режим доступу: <https://eco-city.org.ua/>.

19. Кабінет дослідника «EcoCity». [Електронний ресурс] – Режим доступу: <https://archive.eco-city.org.ua/>.

20. «SaveEcoBot». [Електронний ресурс] – Режим доступу: <https://www.saveecobot.com/>.

21. European Air Quality Index». [Електронний ресурс] – Режим доступу: <https://airindex.eea.europa.eu/AQI/index.html>.

22. Теорія керування і прогнозування у складних системах : підручник / В. Д. Романенко, Ю. Л. Мілявський. – К. : КПІ ім. Ігоря Сікорського, Вид-во «Політехніка», 2024. – 404 с.

23. Машинне навчання : навчальний посібник, призначений для студентів, що навчаються за першим (бакалаврським) рівнем вищої освіти за спеціальностями галузі знань 12 «Інформаційні технології» / Басюк Т. М., Литвин В. В., Захарія Л. М., Кунанець Н. Е., Львів : Видавництво «Новий Світ - 2000», 2019. 335 с.

24. Директива 2008/50/ЄС Європейського Парламенту та Ради від 21 травня 2008 року про якість атмосферного повітря та чистіше повітря для Європи. [https://zakon.rada.gov.ua/laws/show/994_950#Text].

25. Д. О. Шмундяк і В. Б. Мокін, «Метод ідентифікації параметрів гармонік та аномалій періодичного часового ряду на основі адаптивної декомпозиції», Вісник Вінницького політехнічного інституту, вип. 6, с. 46–56, Груд. 2023.

26. Omar, Salima & Ngadi, Md & Jebur, Hamid & Benqdara, Salima. (2013). Machine Learning Techniques for Anomaly Detection: An Overview. *International Journal of Computer Applications*. 79. 10.5120/13715-1478.
27. Julien Lesouple, Cédric Baudoin, Marc Spigai, Jean-Yves Tourneret, Generalized isolation forest for anomaly detection, *Pattern Recognition Letters*, Volume 149, 2021, Pages 109-119, ISSN 0167-8655, <https://doi.org/10.1016/j.patrec.2021.05.022>.
28. Sklearn. API Reference – Available: <https://scikit-learn.org/stable/modules/classes.html>. Accessed: 07.12.2023.
29. Yumin Chen, Duoqian Miao, Hongyun Zhang, Neighborhood outlier detection, *Expert Systems with Applications*, Volume 37, Issue 12, 2010, Pages 8745-8749, ISSN 0957-4174, <https://doi.org/10.1016/j.eswa.2010.06.040>.
30. Markus M. Breunig, Hans-Peter Kriegel, Raymond T. Ng, and Jörg Sander. 2000. LOF: identifying density-based local outliers. *SIGMOD Rec.* 29, 2 (June 2000), 93–104. <https://doi.org/10.1145/335191.335388>.
31. VIEIRA, Rafael G.; LEONE FILHO, Marcos A.; SEMOLINI, Robinson. An Enhanced Seasonal-Hybrid ESD Technique for Robust Anomaly Detection on Time Series. In: *SIMPÓSIO BRASILEIRO DE REDES DE COMPUTADORES E SISTEMAS DISTRIBUÍDOS (SBRC)*, 36. , 2018, Campos do Jordão. Anais. Porto Alegre: Sociedade Brasileira de Computação, 2018 . p. 281-294. ISSN 2177-9384. <https://doi.org/10.5753/sbrc.2018.2422>.
32. Аналіз можливостей Python-бібліотек щодо виявлення аномальних даних у задачі прогнозування стану атмосферного повітря [Електронний ресурс] / Д. О. Шмундяк, Н. С. Іжаковська, Д. О. Литвиненко, А. О. Судець // Матеріали ЛП Науково-технічної конференції підрозділів ВНТУ, Вінниця, 21-23 червня 2023 р. – Електрон. текст. дані. – 2023. – Режим доступу: <https://conferences.vntu.edu.ua/index.php/all-fksa/all-fksa-2023/paper/view/18957>.
33. В. Б. Мокін, О. В. Слободянюк, О. М. Давидюк, і Д. О. Шмундяк, «Інформаційна технологія пошуку можливих джерел підвищеного забруднення річки з використанням моделі Prophet», *Вісник Вінницького політехнічного інституту*, вип. 4, с. 15–24, Верес. 2020.
34. Метод пошуку можливих джерел забруднення атмосферного повітря на основі аналізу аномалій / Д. О. Шмундяк // Матеріали XXIII Міжнародної науково-практичної конференції Математичне моделювання та інформаційно-комунікаційні технології для перемоги та відновлення, Вінниця, 16-17 жовтня 2024 р.

35. Saaty T. L. How to make a decision: The Analytic Hierarchy Process, European Journal of Operational Research, 1990. – 48. – Pages 9-26.

36. Saaty T. L. Multicriteria Decision Making: The Analytic Hierarchy Process. – Pittsburgh: RWS Publ., 1996. – 287 p.

37. Mokin B. I. Decision Support System For The Use Of Funds Received From Higher Education Institution Paid Services / Mokin, Borys I; Mokin, Vitalii B; Mokina, Yuliya V. Actual Problems in Economics, # 3(177), 2016, Pages 372-383.

38. В. Б. Мокін, Є. М. Крижановський, А. Р. Ящолт, і Д. О. Шмундяк, «Технологія проектування мережі спостережень якості атмосферного повітря регіону на основі методу аналізу ієрархій», Наукові праці ВНТУ, вип. 4, Лют 2022.

39. Оптимізація моніторингової мережі стану атмосферного повітря Вінницької зони за даними 2016-2020 рр. [Електронний ресурс] / В. Б. Мокін, Є. М. Крижановський, А. Р. Ящолт, Д. О. Шмундяк, Б. В. Паєвський // Матеріали LI Науково-технічної конференції підрозділів ВНТУ, Вінниця, 31 травня-1 червня 2022 р. – Електрон. текст. дані. – 2023. – Режим доступу: <https://conferences.vntu.edu.ua/index.php/all-fksa/all-fksa-2022/paper/view/16118>.

40. Дослідження впливу параметрів моделі Prophet на точність прогнозування якості атмосферного повітря [Електронний ресурс] / Д. О. Шмундяк // Матеріали LII Науково-технічної конференції підрозділів ВНТУ, Вінниця, 21-23 червня 2023 р. – Електрон. текст. дані. – 2023. – Режим доступу: <https://conferences.vntu.edu.ua/index.php/all-fksa/all-fksa-2023/paper/view/18210>.

41. Огляд підходів до визначення порядку Фур'є у моделі Facebook Prophet для моделювання сезонної складової часового ряду [Електронний ресурс] / Д. О. Шмундяк, А. Лосенко, В. Б. Мокін // Матеріали LII Науково-технічної конференції підрозділів ВНТУ, Вінниця, 21-23 червня 2023 р. – Електрон. текст. дані. – 2023. – Режим доступу: <https://conferences.vntu.edu.ua/index.php/all-fksa/all-fksa-2023/paper/view/17200>.

42. Vitalii Mokin, and Arsen Losenko, “COVID-19 Ukraine daily cases - EDA” Kaggle Notebook. [Електронний ресурс]. Режим доступу: <https://www.kaggle.com/code/vbmokin/covid-19-ukraine-daily-cases-eda>.

43. Dmytro Shmundiak and Vitalii Mokin, “Adaptive decomposition for harmonics and anomalies” Kaggle Notebook. [Electronic resource]. Available: <https://www.kaggle.com/code/dimashmundiak/adaptive-decomposition-for-harmonics-and-anomalies>. Accessed:20.12.2023.

44. Д. О. Шмундяк і В. Є. Копняк, «Метод ідентифікації локальних аномалій значень показників стану довкілля з використанням декомпозиції на півхвилі», Вісник Вінницького політехнічного інституту, вип. 1, с. 88–100, Лют. 2024.

45. Victoria Rodinkova, Serhii Yuriev, Vitalii Mokin, Mariia Kryvopustova, Dmytro Shmundiak, Mykyta Bortnyk, Yevhenii Kryzhanovskyi, Andrii Kurchenko, Bayesian analysis suggests independent development of sensitization to different fungal allergens, World Allergy Organization Journal, Volume 17, Issue 5, 2024, 100908, ISSN 1939-4551, <https://doi.org/10.1016/j.waojou.2024.100908>.

46. Dmytro Shmundiak, “Bayes for air monitoring data” Kaggle Notebook. [Електронний ресурс]. Режим доступу: <https://www.kaggle.com/code/dimashmundiak/bayes-for-air-monitoring-data>.

ДОДАТОК А

ГРОМАДСЬКІ ПРОЄКТИ МОНІТОРИНГУ ЯКОСТІ АТМОСФЕРНОГО ПОВІТРЯ

Важливим суб'єктом моніторингу атмосферного повітря є мережі громадського моніторингу. Прикладом такої мережі є проєкт «EcoCity», який розроблений і функціонує в рамках програми «Чисте повітря для України». Проєкт займається покращенням доступ населення до екологічної інформації. Наразі мережа нараховує понад 250 станцій по всій країні. Кожна з станцій оцінюють забруднення повітря по певному набору основних показників (кількість показників залежить від комплектації станції) і передають дані онлайн у режимі реального часу. Одна з таких станцій розміщена на території ВНТУ і як зазначено раніше передає дані на сервери EcoCity, що дозволяє в режимі реального часу бачити значення усіх доступних показників. Головна сторінка веб-порталу EcoCity і безпосередньо показники забруднення, отримані з станції, розміщеної на території ВНТУ, зображено на рисунку А.1.

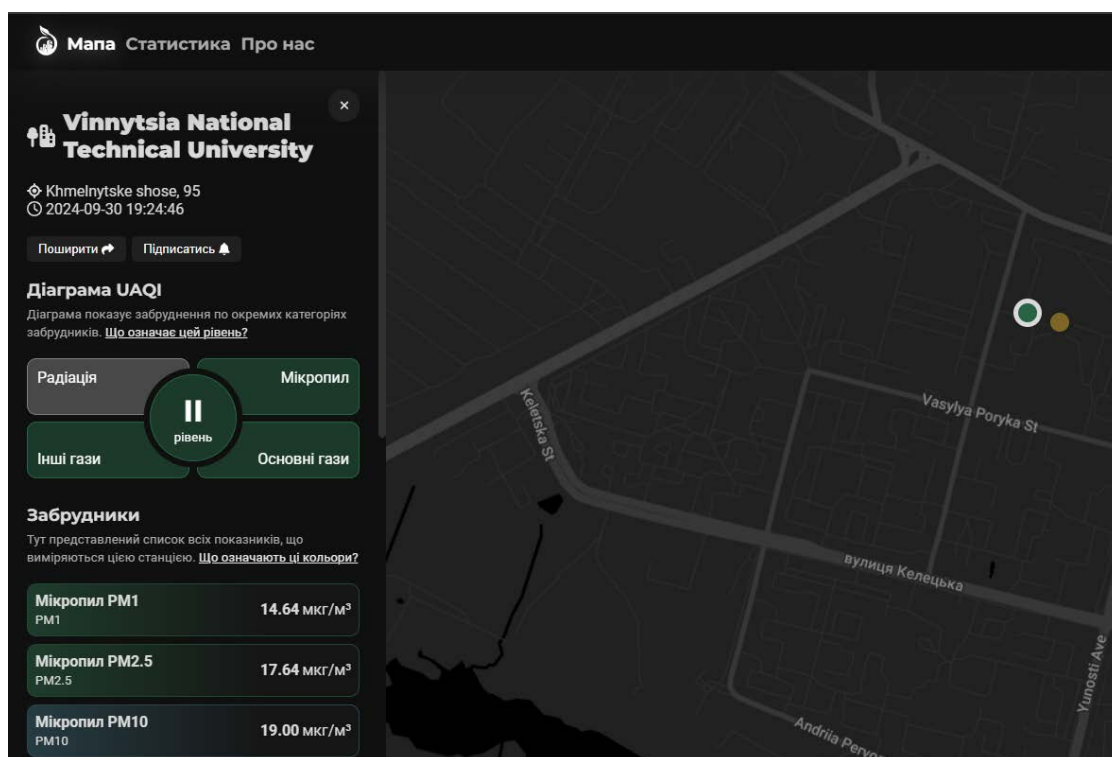


Рисунок А.1 – Головна сторінка веб-сайту EcoCity

Мінімальна комплектація станції дозволяє збирати показники температури, атмосферного тиску, вологості повітря, концентрацію твердих

частинок діаметром до 2.5 мікрметра (PM2.5) та діаметром до 10 мікрметрів (PM10). Загальний перелік можливих показників також включає чадний газ (CO), діоксид азоту (NO2), озон (O3), формальдегід (CH2O), вуглекислий газ (CO2), радіаційний фон, тощо. На рисунку А.2 наведено приклад сторінки з деталями станції моніторингу.

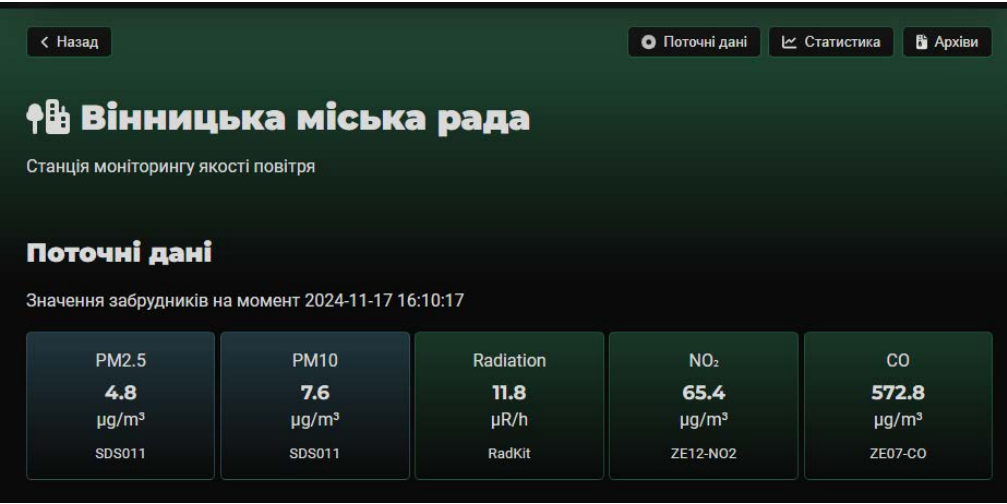


Рисунок А.2 – Головна сторінка веб-сайту EcoCity

На сторінці з деталями, також можна переглядати базову статистику по кожному показнику за останні 7 та останні 30 днів. Приклад даного вікна наведено на рисунку А.3.

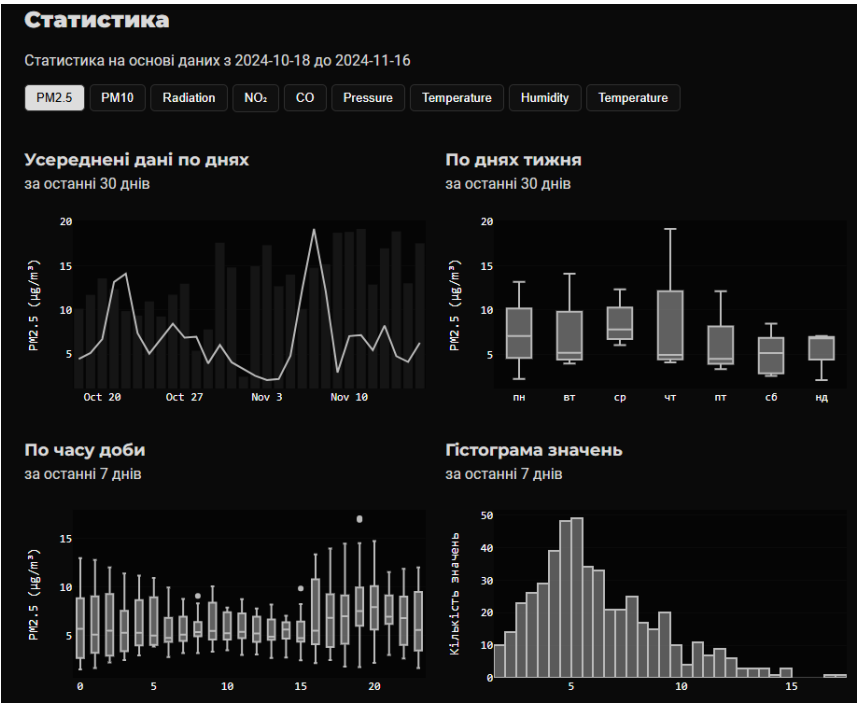


Рисунок А.3 – Сторінка з статистикою по області від веб-сайту EcoCity

Іншим проектом громадського моніторингу є «SaveEcoBot» – проект від громадської організації «SaveDnipro», яка займається аналізом екологічної ситуації в Україні. Проект надає чат-боти у Telegram, Viber та Facebook Messenger і мапу забруднення атмосферного повітря. Проект надає як свої дані, так і агрегує дані від партнерів.

Стан атмосферного повітря можна переглядати у вигляді інтерактивної мапи, яка доступна на веб-сторінці сервісу, яку зображено на рисунку А.4.

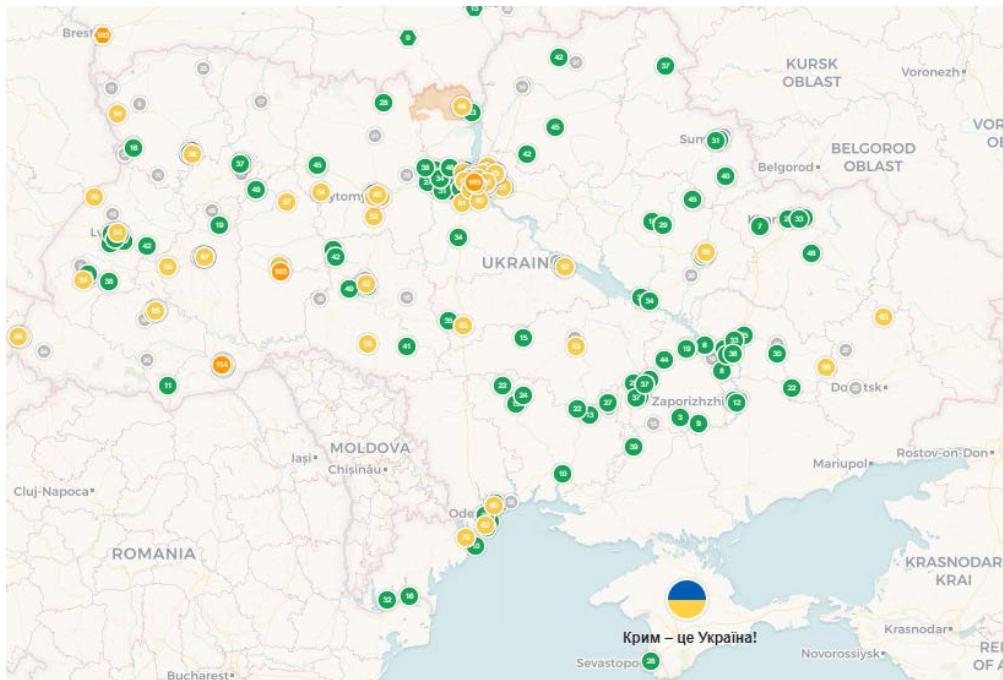


Рисунок А.4 – Сторінка веб-сайту SaveEcoBot

Приклад однієї зі станцій наведено на рисунку А.5

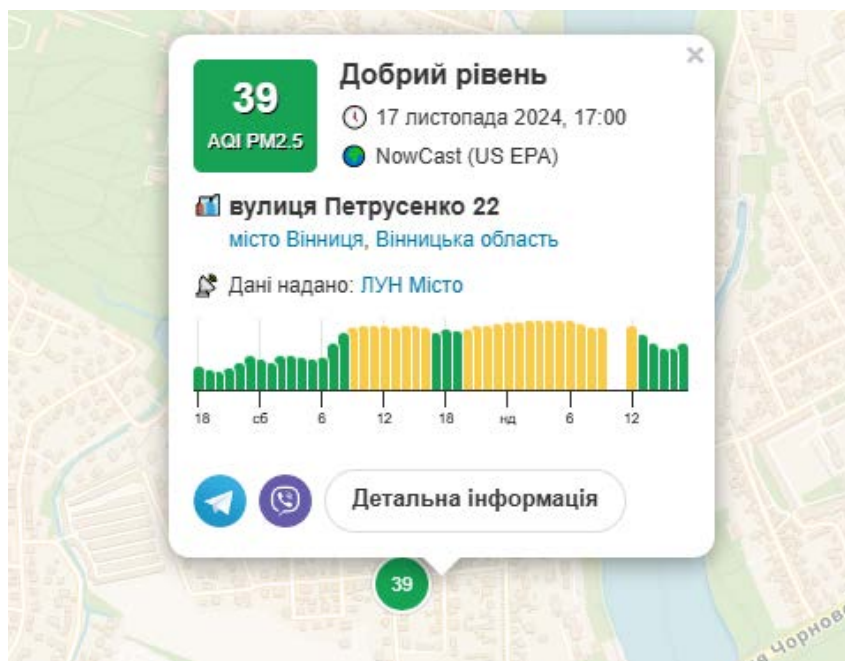


Рисунок А.5 – Сторінка веб-сайту SaveEcoBot з даними з однієї станції

Якість атмосферного повітря представлена у форматі індексу якості повітря. Перейшовши на сторінку деталей по окремій станції можна отримати більш детальні дані та короточасну статистику (рис. А.6, рис А.7, рис А.8).

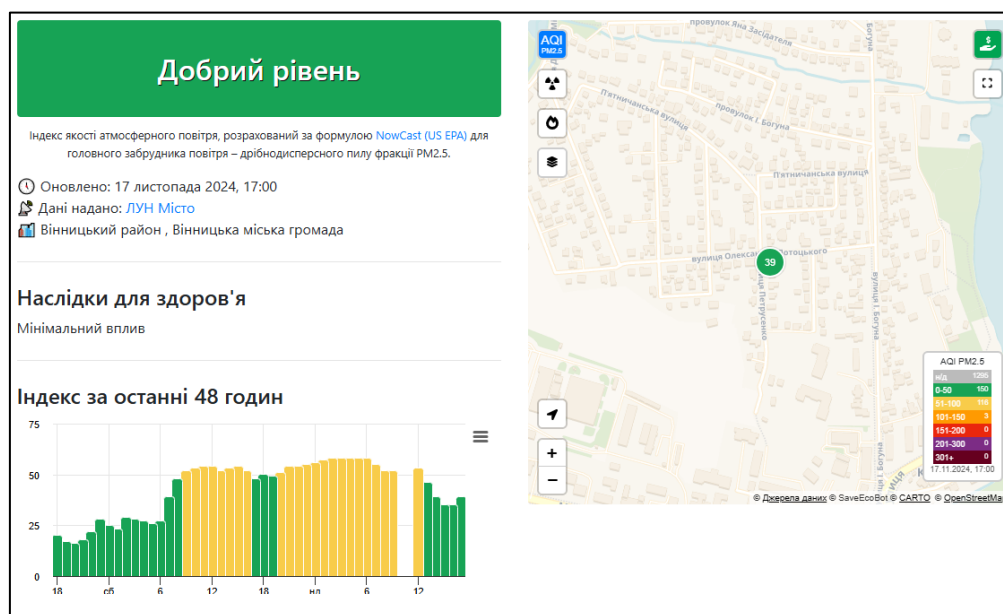
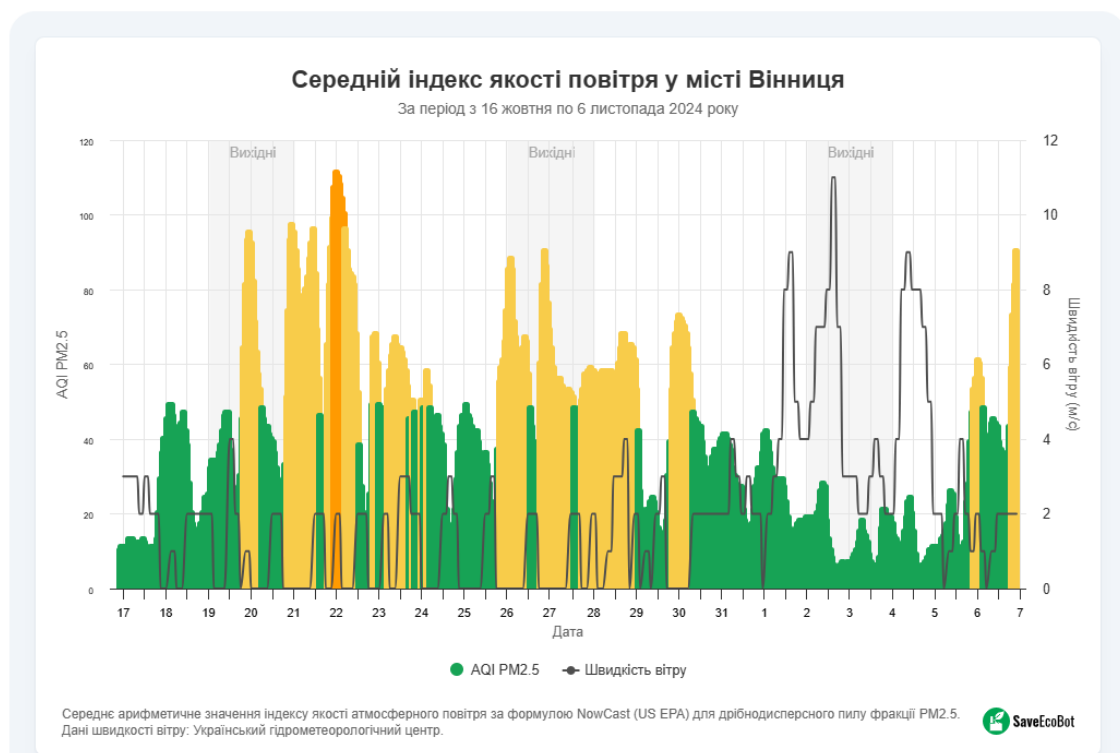
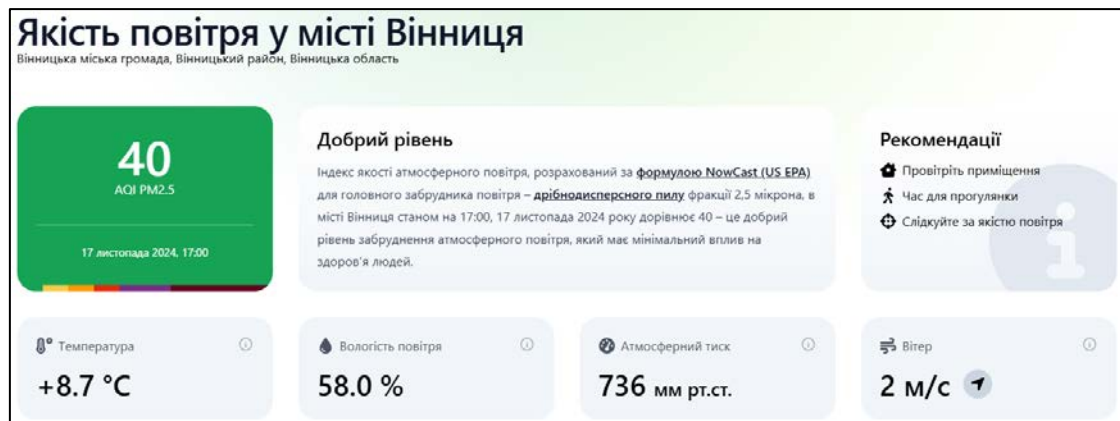


Рисунок А.6 – Сторінка веб-сайту SaveEcoBot з більш детальною інформацією про станцію



ДОДАТОК Б ЄВРОПЕЙСЬКИЙ ІНДЕКС ЯКОСТІ ПОВІТРЯ

Європейський індекс якості повітря (European Air Quality Index) – це стандартизована система, розроблена Європейським агентством з навколишнього середовища (EEA), для інформування громадян країн ЄС щодо якості повітря. Громадяни ЄС можуть зрозуміти якість повітря там, де вони живуть, працюють або подорожують.

Індекс базується на значеннях концентрації п'яти ключових забруднювачів:

- тверді частинки пилу (PM₁₀);
- дрібні частинки пилу (PM_{2,5});
- озон (O₃);
- діоксид азоту (NO₂);
- діоксид сірки (SO₂).

Дана інформація доступна будь якому користувачу на відповідній веб сторінці Європейської агенції довкілля. Приклад цієї сторінки наведено на рисунку Б.1.

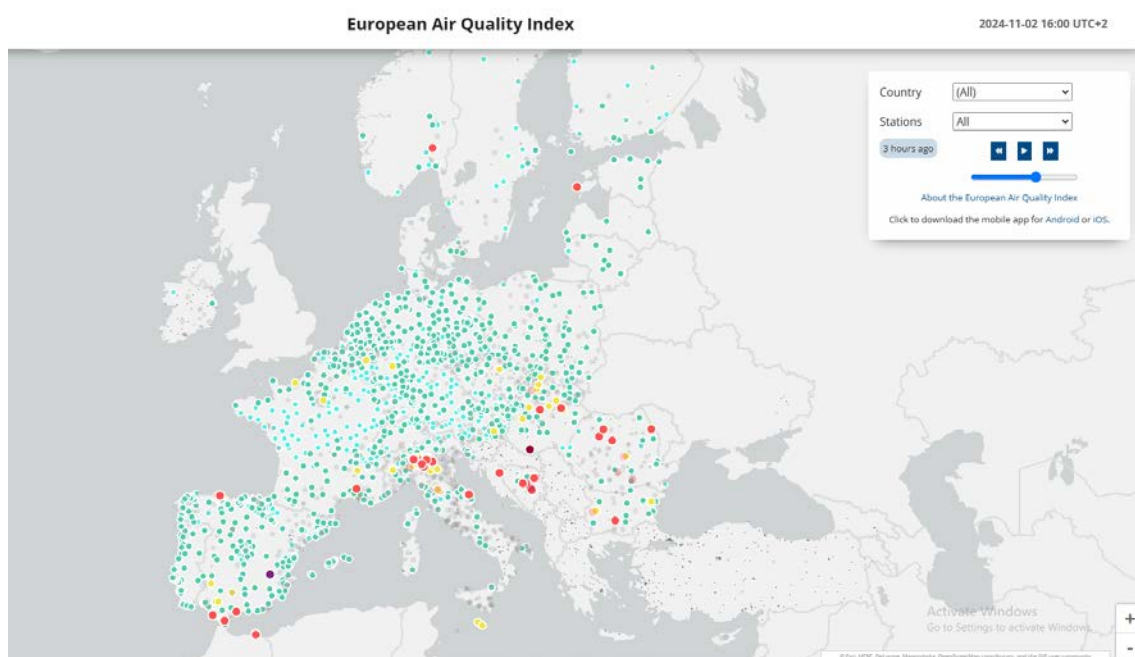


Рисунок Б.1 – Інтерактивна мапа рівня якості повітря на території ЄС

Значення кожного з показників якості повітря за таблицею порівняння перетворюють на відповідну категорію рівня якості, а загальний індекс вираховується по найгіршому рівню серед наявних показників. Таблиця з рівнями якості повітря для кожного з показників наведено на рисунку Б.2.

Pollutant	Index level (based on pollutant concentrations in $\mu\text{g}/\text{m}^3$)					
	Good	Fair	Moderate	Poor	Very poor	Extremely poor
Particles less than $2.5\ \mu\text{m}$ ($\text{PM}_{2.5}$)	0-10	10-20	20-25	25-50	50-75	75-800
Particles less than $10\ \mu\text{m}$ (PM_{10})	0-20	20-40	40-50	50-100	100-150	150-1200
Nitrogen dioxide (NO_2)	0-40	40-90	90-120	120-230	230-340	340-1000
Ozone (O_3)	0-50	50-100	100-130	130-240	240-380	380-800
Sulphur dioxide (SO_2)	0-100	100-200	200-350	350-500	500-750	750-1250

Рисунок Б.2 – Рівні якості повітря та відповідні граничні показники

ДОДАТОК В

УКРАЇНСЬКИЙ ІНДЕКС ЯКОСТІ ПОВІТРЯ

Таблиця В.1 – Український індекс якості повітря (UA AQI)

Категорія	Значення	Для уразливих та чутливих груп
Категорія I	Добра якість повітря	Вміст у повітрі забруднювальних речовин у межах норми та не становить небезпеки. Плануйте діяльність, відпочинок або інші активності на відкритому повітрі без додаткових обмежень, застережень або рекомендацій.
Категорія II	Задовільна якість повітря	Вміст у повітрі забруднювальних речовин у межах норми та не становить небезпеки. Плануйте діяльність, відпочинок або інші активності на відкритому повітрі без додаткових обмежень, застережень або рекомендацій. У рідких випадках дуже невеликої кількості населення, які надзвичайно чутливі до забруднення повітря, можуть виникнути дискомфорт та рефлекторні реакції – кашель, задишка тощо. Надзвичайно чутливим до забруднення повітря людям рекомендуємо зменшити тривалі або важкі фізичні навантаження на відкритому повітрі.
Категорія III	Якість повітря несприятлива	Вміст у повітрі забруднювальних речовин у межах норми, проте уразливі та чутливі до забруднення категорії населення можуть мати наслідки для стану здоров'я при довготривалому впливі. Врахуйте спеціальні обмеження, застереження та рекомендації під час планування діяльності, відпочинку або інших активностей на відкритому повітрі. Рекомендуємо уразливим та чутливим до забруднення категоріям населення зменшити тривалі або важкі фізичні навантаження на відкритому повітрі. Якщо ви відчуваєте дискомфорт під час дихання та прояви інших рефлекторних реакцій – рекомендуємо дотримуватися звичайних порад та плану лікування від вашого лікаря. Люди хворі на астму, ХОЗЛ та респіраторні захворювання можуть відчувати посилення звичних симптомів та рефлекторних реакцій. Рекомендуємо сумлінно дотримуватися свого плану терапії хронічних захворювань та прийому ліків, які призначив ваш лікар. Люди із хронічними захворюваннями серцево-судинної системи можуть відчувати додаткові та посилені рефлекторні реакції – підвищене серцебиття, задишка або незвичайна втома. У випадку довготривалого прояву цих реакцій рекомендуємо звернутися за консультацією до вашого лікаря та сумлінно виконувати його настанови.

Продовження таблиці В.1

Категорія	Значення	Для уразливих та чутливих груп
Категорія IV	Погана якість повітря	Вміст у повітрі забруднювальних речовин становить небезпеку при довготривалому впливі. Усі категорії населення можуть відчувати загострення рефлекторних реакцій та мати наслідки для здоров'я при довготривалому впливі. Діють спеціальні обмеження, застереження та рекомендації для тривалості діяльності, відпочинку або інших активностей на відкритому повітрі. Усім уразливим та чутливим до забруднення категоріям населення рекомендуємо виключити та перенести на інших час тривалі або важкі фізичні навантаження на відкритому повітрі. Чутливі до забруднення категорії населення, які мають хронічні захворювання, можуть відчувати значне посилення звичних симптомів та рефлекторних реакцій. Рекомендуємо сумлінно дотримуватися свого плану терапії хронічних захворювань та прийому ліків, які призначив ваш лікар. У випадку довготривалого прояву рефлекторних реакцій рекомендуємо звернутися за консультацією до вашого лікаря та сумлінно виконувати його настанови.
Категорія V	Дуже погана якість повітря	Для уразливих та чутливих до забруднення категорій населення вміст у повітрі забруднювальних речовин становить небезпеку навіть при короткостроковому впливі. Усі категорії населення можуть відчувати сильне загострення рефлекторних реакцій та мати наслідки для здоров'я небезпеку навіть при короткостроковому впливі. Для уразливих та чутливих до забруднення категорій населення діють спеціальні обмеження, застереження та рекомендації для перебування на відкритому повітрі для будь-якої діяльності. Рекомендуємо виключити або перенести на інший час будь-яку діяльність на відкритому повітрі. Якщо у вас виникає рефлекторна реакція на забруднене повітря – кашель, задишка, подразнення слизових оболонок носоглотки, біль в очах тощо – рекомендуємо використовувати засоби індивідуального захисту органів дихання. У випадку довготривалого прояву рефлекторних реакцій або інших відчутних наслідків для вашого здоров'я рекомендуємо звернутися за консультацією до вашого лікаря та сумлінно виконувати його настанови.

Продовження таблиці В.1

Категорія	Значення	Для уразливих та чутливих груп
Категорія VI	Надзвичайно погана якість повітря	Вміст у повітрі забруднювальних речовин становить небезпеку навіть при короткостроковому впливі. Усі категорії населення можуть відчувати сильне загострення рефлексорних реакцій та мати наслідки для здоров'я небезпеку навіть при короткостроковому впливі. Для всіх категорій населення діють спеціальні обмеження, застереження та рекомендації для перебування на відкритому повітрі для будь-якої діяльності. Рекомендуємо виключити або перенести на інший час будь-яку діяльність на відкритому повітрі. Якщо у вас виникає рефлексорна реакція на забруднене повітря – кашель, задишка, подразнення слизових оболонок носоглотки, біль в очах тощо – рекомендуємо використовувати засоби індивідуального захисту органів дихання. У випадку довготривалого прояву рефлексорних реакцій або інших відчутних наслідків для вашого здоров'я рекомендуємо звернутися за консультацією до вашого лікаря та сумлінно виконувати його настанови.
Категорія VII	Якість повітря тимчасово невідома, причини технічного характеру	Через різні технічні причини станції моніторингу якості повітря на території, яка вас цікавить, не передають результати спостережень за цим параметром. Якщо під час дихання ви відчуваєте дискомфорт або маєте виражені рефлексорні реакції - рекомендуємо обмежити або перенести на інший час будь-яку діяльність на відкритому повітрі або використовувати засоби індивідуального захисту органів дихання. У випадку довготривалого прояву рефлексорних реакцій або інших відчутних наслідків для вашого здоров'я рекомендуємо звернутися за консультацією до вашого лікаря та сумлінно виконувати його настанови.

ДОДАТОК Г

ФОРМАЛІЗАЦІЯ ІНФОРМАЦІЇ У ВІДПОВІДЯХ ЗАПИТІВ НА ПУБЛІЧНУ ІНФОРМАЦІЮ, ЗБИРАННЯ ЯКОЇ ВИМАГАЄ НАКАЗ МІНДОВКІЛЛЯ №147 ВІД 25.02.21

Таблиця Г.1 – Формалізація інформації у відповідях запитів на публічну інформацію

Код	Вид інформації	Джерело	Доступна інформація	Невизначеність та недоліки
D1	Інформація про кількість населення	Головне управління статистики області	Кількість населення по роках	1. Як правило, інформація доступна тільки за повні роки, наприклад, в кінці 2021 р. вона є тільки за 2020 р.
D2	Інформація про найбільших забруднювачів довкілля	Головне управління статистики області	Назви, адреси підприємств	1. Як правило, інформація доступна тільки за повні роки, наприклад, в кінці 2021 р. вона є тільки за 2020 р., 2. Інформація є тільки по підприємствах, які звітують згідно форми 2-ТП «Повітря».
P1	Координати місць розташування виробничих потужностей підприємств, які є у списку D1	Орган облдержадміністрації, відповідальний за управління якістю атмосферного повітря та Міндовкілля України	Адреси чи GPS-координати центрів усіх майданчиків виробничих потужностей із дозволів на гранично допустимі викиди	1. Частина дозволів видає Міндовкілля, частину – обласні органи і буває складно знайти де саме знаходиться потрібний дозвіл. 2. Часто надаються не координати, а – адреса будівлі, однак, не завжди є впевненість, що це – саме адреса викидів, а не – офісу підприємства, розташованому, як правило, в іншому місці.
S1-01 ... S1-N	Значення N показників якості атмосферного повітря	Обласний центр з гідрометеорології Державної служби України з надзвичайних ситуацій (ДСНС)	Усереднені за добу значення концентрації N показників на невеликій кількості пунктів	1. Документи [1-5] містять вимоги щодо показників PM2.5 і PM10 (тверді частки пилу, розміром 2.5 і 10 мкм, відповідно), а ДСНС та МОЗ, як правило, вимірюють тільки недиференційований пил, який з порогамі PM2.5 і PM10 не порівнюєш. Тому є сенс на початковому етапі враховувати перевищення й гранично допустимих концентрацій (ГДК). 2. Різні суб'єкти моніторингу можуть вимірювати показники за різними методиками, що ускладнює їх порівняння. Крім того, по більшості показників ДСНС усереднює дані, які спостерігає весь час в одному місці, а лабораторії МОЗ, як правило, роблять епізодичні вимірювання, але на великій мережі спостережень.
S2-01-M	Значення M показників якості атмосферного повітря	Обласний лабораторний центр Міністерства охорони здоров'я України	Значення концентрації M показників у багатьох пунктах, але за порівняно невелику кількість діб	
I1	Кількість вразливого населення в регіоні	Головне управління статистики області	Кількість вразливого населення в регіоні по роках.	1. Як правило, інформація доступна тільки за повні роки, наприклад, в кінці 2021 р. вона є тільки за 2020 р.

ДОДАТОК Д

РОЗКРИТТЯ НЕВИЗНАЧЕНОСТІ

Таблиця Д.1 – Розкриття невизначеності запропонованими та удосконаленими методами системного аналізу стану атмосферного повітря

№ п/п	Вид декомпозиції	Вид невизначеності	Підрозділ чи пункт роботи, де було розкрито невизначеність
1.	D_{DPSIR}	U_{S1}	Підрозділ 2.2
2.	D_{DPSIR}	U_{T3}	Пункт 2.3.3
3.	D_{DPSIR}	U_{T5}	Пункт 2.3.3
4.	D_{DPSIR}	U_{T1}	Пункт 2.3.4
5.	D_{DPSIR}	U_{P2}	Підрозділ 3.1
6.	D_{DPSIR}	U_{P1}	Підрозділ 3.1
7.	D_{G-O}	U_{T6}	Підрозділ 3.1
8.	$D_T(\Delta T)$	U_{S2}	Підрозділ 3.2
9.	$D_T(\Delta T)$	U_{ST1}	Підрозділ 3.3
10.	$D_{SEC}(\Theta)$	U_{ST2}	Підрозділ 3.3
11.	$D_{SEC}(\Theta)$	U_{SC1}	Підрозділ 3.4

ДОДАТОК Е

ЗАГАЛЬНА ХАРАКТЕРИСТИКА ЗОНИ «ВІННИЦЬКА»

Зона «Вінницька» розташована на правобережжі Дніпра в межах Придніпровської та Подільської височин. Зона «Вінницька» в межах Вінницької області розташована в центральній лісостеповій частині України, в середній течії річок Дністра та Південного Бугу.

Територія зони «Вінницька» становить 26404,4 км². До адміністративно-територіального складу області входять: 5 районів, 18 міських, 28 селищних рад та 661 сільська рада, в т.ч. 63 об'єднані територіальні громади, 1504 населених пункти, з них 29 селищ міського типу, 18 міст, у тому числі 6 міст обласного значення. Зона «Вінницька» межує з агломерацією «Вінниця», розташованою дещо північніше від центру зони, та має найбільше сусідніх зон серед всіх областей країни. На заході межує з зоною «Чернівецька» та «Хмельницька», на півночі з зоною «Житомирська», на сході – із зонами «Київська», «Кіровоградська» та «Вінницька», на півдні – із зоною «Одеська» та з Республікою Молдова (202 км державного кордону) (рис. Е.1).

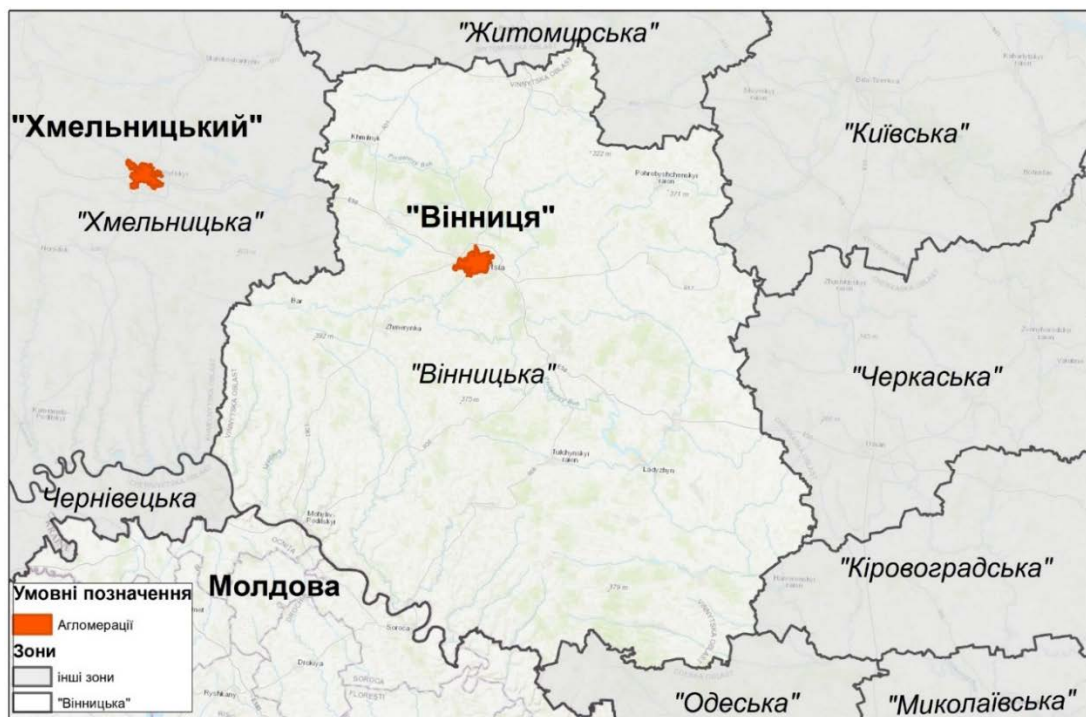


Рисунок Е.1 – Географічне положення зони «Вінницька»

Зона «Вінницька» розміщена в лісостеповій зоні центральної частини Правобережної України. Річкою Південний Буг територія області ділиться на дві частини: лівобережну, яка відноситься до Придніпровської височини і правобережну – до Подільського плато. Поверхня зони

«Вінницька» – підвищене плато, що знижується в напрямі з північного заходу на південний схід. Більша частина території зони «Вінницької» розташована в межах Українського кристалічного щита. Складна геологічна історія території вплинула на формування рельєфу (рис. Е.2).



Рисунок Е.2 – Рельєф зони «Вінницька»

Значний вплив на формування рельєфу також спричинила робота протікаючих вод, розгалужена чисельними долинами річок, ярами та балками, особливо в районі Придністров'я. Територією зони проходить вододіл басейнів річок Південний Буг та Дністер. У центральній частині області з північно-західного на південно-східний напрямок протікає р. Південний Буг, по південно-західній межі області тече р. Дністер. На території зони протікають понад 200 річок завдовжки понад 10 км кожна. Вони належать до басейнів Південного Бугу (Згар, Рів, Дохна, Соб, Савранка), Дністра (Мурафа, Лядова, Марківка, Русава, Немія) та Дніпра (Рось, Гнилоп'ять, Гуйва). Пересічна густота річкової мережі становить $0,38 \text{ км/км}^2$. В межах зони – 56 водосховищ, загальною площею водного дзеркала 11167 га; найбільше – Ладизинське водосховище (2,2 тис. га), більше 5700 ставків загальною площею водного дзеркала біля 30,0 тис. га.

У зоні «Вінницька» поширені лісостепові ландшафти. В лісах переважають широколистяні породи дерев: граб, клен, липа, дуб, ясен. Трав'яниста рослинність характеризується великою різноманітністю.

Лише диких рослин нараховується біля тисячі видів. Зона «Вінницька» в геоструктурному плані, в основному, розташована на південно-західній окраїні українського кристалічного масиву (щита), складеного архей-протерозойськими метаморфічними та магматичними породами, вік яких сягає 1,5 - 3,5 мільярда років. Це – так званий, кристалічний фундамент. І тільки південно-західна окраїна зони розташована на Волино-Подільській плиті, де породи фундаменту перекриті відносно потужною товщею осадових відкладів та рідше – вулканічних.

На території зони породи фундаменту – гнейси, кристалічні сланці, мігматити, граніти та більш специфічні утворення – чарнокіти, ендербіти, дайки габбро-діабазів та інші залягають на незначних глибинах – від безпосередніх виходів на поверхню, переважно на схилах та в долинах річок, до 50-100 м на водорозділах, і лише на Наддністрянщині вони занурюються до глибин в 150-300 м і більше. Глибина поширення самих кристалічних порід сягає десятків кілометрів. Ці породи перетерпіли кілька етапів метаморфізму та інтенсивних деформацій. Вони часто зм'яті в складки, в окремих зонах інтенсивно роздроблені, пронизані жилами кварцу та зонами мінералізації іншого складу.

У Придністровській частині зони на породах фундаменту залягає специфічний проміжний верхньопротерозойський комплекс, складений осадовими (пісковики, глинисті сланці) та вулканічними (базальти, туфи) породами. Вік цих утворень оцінюється приблизно в 600-700 мільйонів років. Вони залягають на глибині до 100-150 м на водорозділах, а в долинах річок часто виходять безпосередньо на поверхню. Потужність цієї товщі змінюється від перших метрів на відстані в 30-50 км північніше Дністра і до 100-150 м неподалік його русла. В цих відкладах іноді теж спостерігаються прояви деформацій, слабого метаморфізму та різного типу мінералізації.

Майже на всій території зони верхня частина кристалічних порід фундаменту в мезозойську еру (100-200 мільйонів років), під впливом вологого та теплого клімату зазнала інтенсивних змін (хімічного звітрювання) з утворенням кори звітрювання, переважно каолінового складу. Потужність каолінових кор, що збереглися, місцями сягає до 100 м. На породах фундаменту, чи їхніх корах звітрювання, а на Наддністрянщині на утвореннях проміжного верхньопротерозойського комплексу, залягають осадові відклади платформного чохла – вапняки, глини, піски та інші. При цьому, знову ж таки тільки на Наддністрянщині, цей комплекс починається з відкладів крейдового періоду - крейдоподібних вапняків та мергелів, опок, трепелу, глауконітових пісків з фосфоритами, вік яких сягає близько 100 мільйонів років. Сумарна потужність крейдових порід місцями сягає до 30-80 м. В окремих місцях зони в складі платформеного чохла присутні відклади палеогенового

періоду (40-60 мільйонів років) – піски, вуглисті глини, місцями з лінзами бурого вугілля.

На всій території зони «Вінницької» поширені відклади неогенового періоду (2-25 мільйонів років) – глини, піски та вапняки в її південно-західній частині. Їх потужності сягають до 30-50, на півдні – до 100 м. Завершують розріз чохла утворення плейстоценового (четвертинного) періоду (до 2 мільйонів років) – суглинки на плато та схилах, піщано-гравійно-галечникові наноси на Дністровських терасах, піски уздовж інших річок та льодовикові піщано-глинисті відклади в північній половині області. Потужність цих відкладів, поширених більш ніж на 95% площі області, коливається від перших метрів до 10-20, рідше 30 м.

Клімат Вінницької області помірно-континентальний: помірного та достатнього теплозабезпечення, достатнього зволоження, лише у Придністров'ї недостатнього зволоження. За своїм географічним розташуванням територія зони знаходиться у сфері впливу насичених вологою атлантичних повітряних мас та периферійної частини сибірського (азійського) антициклону, для якого характерні сухі холодні континентальні повітряні маси. На клімат впливають також повітряні маси з Арктики та Середземномор'я.

В літню пору переважають вологі вітри західного та північно-західного румбів, найбільший їх вплив спостерігається на північний захід від лінії Могилів-Подільський – Гайсин. В холодну пору (жовтень – квітень) відчутний вплив (особливо на південний схід від цієї лінії) сибірського антициклону з вітрами південних та південно-східних румбів. Найхолодніший місяць по всій зоні – січень, найтепліший – липень. Середні амплітуди коливань температури протягом року не перевищують 25⁰С. Під впливом континентальних повітряних мас іноді спостерігається зниження температури в окремі дні до -32...-38⁰С.

Влітку підвищення до +37⁰С, найвищі температури спостерігаються у липні-серпні. Середньорічні суми опадів на території області складають 440-590 мм. Найбільша кількість опадів буває на північному заході території Вінниччини. Максимум опадів припадає на травень – липень (130-170 мм). Найменш вологими є зимові місяці, на холодну пору року припадає 25% опадів: у грудні-лютому випадає 65-80 мм опадів. Перехід від однієї пори року до іншої відбувається поступово. Стійкий перехід добової температури через 0⁰С є початком весни та відбувається найчастіше у другій декаді березня. Весна триває близько двох місяців. Характерними особливостями весни є інтенсивне підвищення денної температури, сходить стійкий сніговий покрив та відтає ґрунт. Літо триває з другої половини травня до першої половини вересня, денні температури становлять у травні +18... +20⁰С, у липні

+21...+25⁰C. В цей же час випадає найбільше опадів, переважно у вигляді злив. Кількість днів з опадами поступово зменшується з наближенням осені. Осінь починається з переходом середньодобової температури через +10⁰C в бік зниження. Настання осені (перша декада жовтня) супроводжується заморозками, загальним зниженням температури, зменшенням кількості опадів. Характерною особливістю осені на Вінниччині є повернення теплих сонячних днів. Осінь закінчується в кінці листопада, коли середньодобові температури переходять через 0⁰C в бік мінусових температур. До початку зими середньодобові температури всюди нижче 0⁰C, але вище -5⁰C, погода нестійка: морозні дні змінюються відлигами, не раз утворюється та сходить сніговий покрив. Відлиги характерні і впродовж зими, температура повітря інколи підвищується до +10...+13⁰C.

ДОДАТОК Є
АНАЛІЗ ОСНОВНИХ ДЖЕРЕЛ ЗАБРУДНЕННЯ
АТМОСФЕРНОГО ПОВІТРЯ ЗОНИ «ВІННИЦЬКА»

Зібрано та систематизовано дані за 2016-2020 рр. про динаміку зміни кількості джерел забруднення атмосферного повітря на території досліджуваної зони (табл. Є.1).

Таблиця Є.1 – Джерела забруднення атмосферного повітря

		2016 рік	2017 рік	2018 рік	2019 рік	2020 рік
1	Загальна кількість підприємств, що здійснюють викиди забруднювальних речовин в атмосферне повітря, од	958	1205	1539	1785	2031
2	Загальна кількість (одиниць) діючих дозволів на викиди забруднювальних речовин в атмосферне повітря, об'єкт якого належить до:	407	523	685	343	272
	першої групи	9	11	15	15	15
	другої групи та третьої групи	407	523	685	343	272
3	Кількість зареєстрованих транспортних засобів, тис. од, з них такі, що належать:					
	юридичним особам, од					
	фізичним особам, од.					
4	Протяжність автомобільних доріг, тис. км				7,520	7,520
	з них з твердим покриттям				7,017	7,017
5	Інші джерела забруднення, од					
	кількість аеропортів	1	1	1	1	1
	кількість морських/річкових портів					

Продовження таблиці Є.1

		2016 рік	2017 рік	2018 рік	2019 рік	2020 рік
	кількість об'єктів утворення, оброблення та утилізації відходів, місць видалення відходів	247 759	260 759	266 761	272 763	275 763
6	Природні джерела (за наявності)					

Також, проаналізовано інформацію про забруднення атмосферного повітря досліджуваної зони. Результати аналізу наведено в таблиці Є.2.

Таблиця Є.2 – Інформація про забруднення атмосферного повітря

		2016 рік	2017 рік	2018 рік	2019 рік	2020 рік
1	Загальний обсяг викидів забруднювальних речовин, тис. т	107807,5	142217,7	94906,9	96895,14	75578,36
2	Викиди забруднювальних речовин від стаціонарних джерел (тис. т) всього.	107753,9	142159,1	94841	96837,040	75514,357
3	Викиди забруднювальних речовин від пересувних джерел, тис. т	53,6	58,6	65,9	58,1	64,0

Використовуючи дані геопросторової прив'язки, здійснено геокодування стаціонарних джерел забруднення досліджуваної зони. На рисунку Є.1 приведено карту основних стаціонарних джерел забруднення зони «Вінницька».

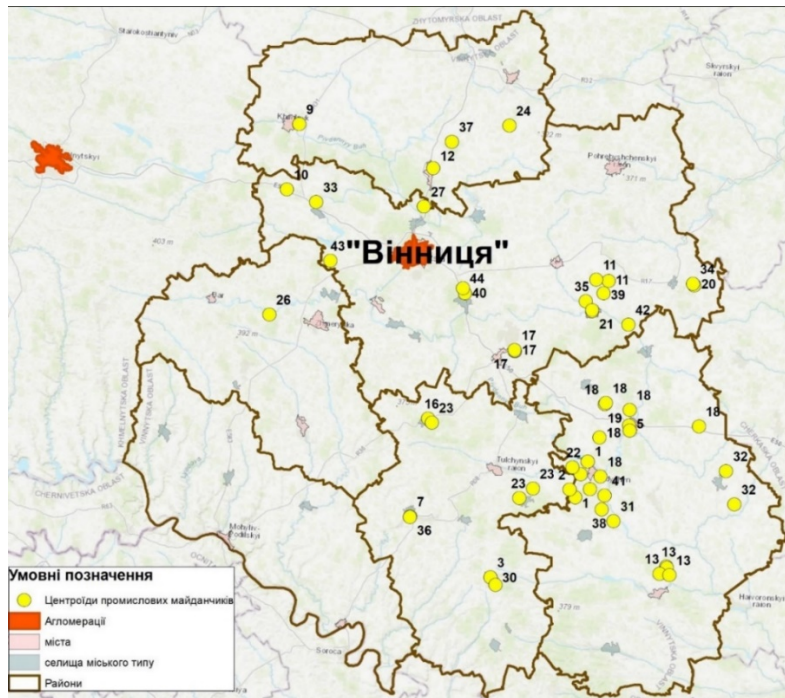


Рисунок Є.1 – Основні стаціонарні джерела забруднення зони «Вінницька»

ДОДАТОК Ж

АНАЛІЗ ЯКОСТІ АТМОСФЕРНОГО ПОВІТРЯ ТА ВИБІР РЕЖИМІВ СПОСТЕРЕЖЕНЬ

Таблиця Ж.1 -Попередня оцінка просторового розподілу концентрацій забруднювальних речовин

№	Територія роз-ташування	Забруднювальна речовина	Метод оцінки	Опис методу оцінки	Джерело даних та інформації, що використовувались
1	Міська	Діоксид сірки	інвентаризація викидів, моделювання, об'єктивне оцінювання	Звіт про результати попереднього оцінювання (додаток 4 до Програми)	Дані Головного управління статистики в Вінницькій області; Моніторингові спостереження Вінницького обласного лабораторного центру Міністерства охорони здоров'я України; Інформація Національного центру управління та випробувань космічних засобів Державного космічного агентства України.
2	Міська	Діоксид азоту	інвентаризація викидів, моделювання, об'єктивне оцінювання	Звіт про результати попереднього оцінювання (додаток 4 до Програми)	Дані Головного управління статистики в Вінницькій області; Моніторингові спостереження Вінницького обласного лабораторного центру Міністерства охорони здоров'я України; Інформація Національного центру управління та випробувань космічних засобів Державного космічного агентства України.
3	Міська	Оксид вуглецю	інвентаризація викидів, моделювання, об'єктивне оцінювання	Звіт про результати попереднього оцінювання (додаток 4 до Програми)	Дані Головного управління статистики в Вінницькій області; Моніторингові спостереження Вінницького обласного лабораторного центру Міністерства охорони здоров'я України; Інформація Національного центру управління та випробувань космічних засобів Державного космічного агентства України.

Продовження таблиці Ж.1

№	Територія роз-ташування	Забруднювальна речовина	Метод оцінки	Опис методу оцінки	Джерело даних та інформації, що використовувались
4	Міська	Пил	інвентаризація викидів, моделювання, об'єктивне оцінювання	Звіт про результати попереднього оцінювання (додаток 4 до Програми)	Дані Головного управління статистики в Вінницькій області; Моніторингові спостереження Вінницького обласного лабораторного центру Міністерства охорони здоров'я України; Інформація Національного центру управління та випробувань космічних засобів Державного космічного агентства України.
5	Міська	Формальдегід	інвентаризація викидів, моделювання, об'єктивне оцінювання	Звіт про результати попереднього оцінювання (додаток 4 до Програми)	Дані Головного управління статистики в Вінницькій області; Моніторингові спостереження Вінницького обласного лабораторного центру Міністерства охорони здоров'я України; Інформація Національного центру управління та випробувань космічних засобів Державного космічного агентства України.
6	Міська	Аміак	інвентаризація викидів, моделювання, об'єктивне оцінювання	Звіт про результати попереднього оцінювання (додаток 4 до Програми)	Дані Головного управління статистики в Вінницькій області; Моніторингові спостереження Вінницького обласного лабораторного центру Міністерства охорони здоров'я України; Інформація Національного центру управління та випробувань космічних засобів Державного космічного агентства України.

Продовження таблиці Ж.1

№	Територія роз-ташування	Забруднювальна речовина	Метод оцінки	Опис методу оцінки	Джерело даних та інформації, що використовувались
7	Міська	Озон	моделювання, об'єктивне оцінювання	Звіт про результати попереднього оцінювання (додаток 4 до Програми)	Інформація Національного центру управління та випробувань космічних засобів Державного космічного агентства України.
8	Сільська	Діоксид сірки	інвентаризація викидів, моделювання, об'єктивне оцінювання	Звіт про результати попереднього оцінювання (додаток 4 до Програми)	Дані Головного управління статистики в Вінницькій області; Моніторингові спостереження Вінницького обласного лабораторного центру Міністерства охорони здоров'я України; Інформація Національного центру управління та випробувань космічних засобів Державного космічного агентства України.
9	Сільська	Діоксид азоту	інвентаризація викидів, моделювання, об'єктивне оцінювання	Звіт про результати попереднього оцінювання (додаток 4 до Програми)	Дані Головного управління статистики в Вінницькій області; Моніторингові спостереження Вінницького обласного лабораторного центру Міністерства охорони здоров'я України; Інформація Національного центру управління та випробувань космічних засобів Державного космічного агентства України.
10	Сільська	Оксид вуглецю	інвентаризація викидів, моделювання, об'єктивне оцінювання	Звіт про результати попереднього оцінювання (додаток 4 до Програми)	Дані Головного управління статистики в Вінницькій області; Моніторингові спостереження Вінницького обласного лабораторного центру Міністерства охорони здоров'я України; Інформація Національного центру управління та випробувань космічних засобів Державного космічного агентства України.

Продовження таблиці Ж.1

№	Територія роз- ташування	Забруднюва- льна речовина	Метод оцінки	Опис методу оцінки	Джерело даних та інформації, що вико- ристовувались
11	Сільська	Пил	інвентаризація викидів, моделювання, об'єктивне оцінювання	Звіт про результати попереднього оціню- вання (додаток 4 до Програми)	Дані Головного управління статистики в Вінницькій області; Моніторингові спостереження Вінницького обласного лабораторного центру Міністерства охорони здоров'я України; Інформація Національного центру управління та випробувань космічних засобів Держа- вного космічного агентства України.
12	Сільська	Формальдегід	інвентаризація викидів, моделювання, об'єктивне оцінювання	Звіт про результати попереднього оціню- вання (додаток 4 до Програми)	Дані Головного управління статистики в Вінницькій області; Моніторингові спостереження Вінницького обласного лабораторного центру Міністерства охорони здоров'я України; Інформація Національного центру управління та випробувань космічних засобів Держа- вного космічного агентства України.
13	Сільська	Аміак	інвентаризація викидів, моделювання, об'єктивне оцінювання	Звіт про результати попереднього оціню- вання (додаток 4 до Програми)	Дані Головного управління статистики в Вінницькій області; Моніторингові спостереження Вінницького обласного лабораторного центру Міністерства охорони здоров'я України; Інформація Національного центру управління та випробувань космічних засобів Держа- вного космічного агентства України.
14	Сільська	Озон	моделювання, об'єктивне оцінювання	Звіт про результати попереднього оціню- вання (додаток 4 до Програми)	Інформація Національного центру уп- равління та випробувань космічних за- собів Державного космічного агентства України.

Таблиця Ж.2 – Встановлений режим оцінювання в зоні

№	Забруднювальна речовина	Встановлений режим оцінювання	Обґрунтування вибору режиму оцінювання
1	Діоксид сірки	режим фіксованих вимірювань	Згідно моніторингових спостережень Вінницького обласного лабораторного центру Міністерства охорони здоров'я України, рівень діоксиду сірки є вищим верхнього порогу оцінювання.
2	Діоксид азоту	режим фіксованих вимірювань	Згідно моніторингових спостережень Вінницького обласного лабораторного центру Міністерства охорони здоров'я України, рівень діоксиду сірки є вищим верхнього порогу оцінювання.
3	Оксид вуглецю	режим фіксованих вимірювань	Згідно моніторингових спостережень Вінницького обласного лабораторного центру Міністерства охорони здоров'я України, рівень діоксиду сірки є вищим верхнього порогу оцінювання.
4	Тверді частки (ТЧ ₁₀)	не встановлено	На даний час відсутня достатня кількість вимірювань для порівняння з порогами оцінювання та встановлення режимів оцінювання. Для накопичення таких даних у зоні заплановано заходи щодо встановлення пунктів спостережень та вдосконалення наявних мереж спостереження (інформацію про такі заходи наведено у розділі IV Програми).
5	Тверді частки (ТЧ _{2,5})		
6	Озон		
7	бензол		
8	бенз(а)пірен		
9	арсен		
10	свинець		
11	кадмій		
12	нікель		
13	ртуть		

ДОДАТОК 3
ДИНАМІКА ВИКИДІВ ОСНОВНИХ ЗАБРУДНЮВАЛЬНИХ РЕЧОВИН ВІД СТАЦІОНАРНИХ
ДЖЕРЕЛ НА ТЕРИТОРІЇ ЗОНИ «ВІННИЦЬКА»

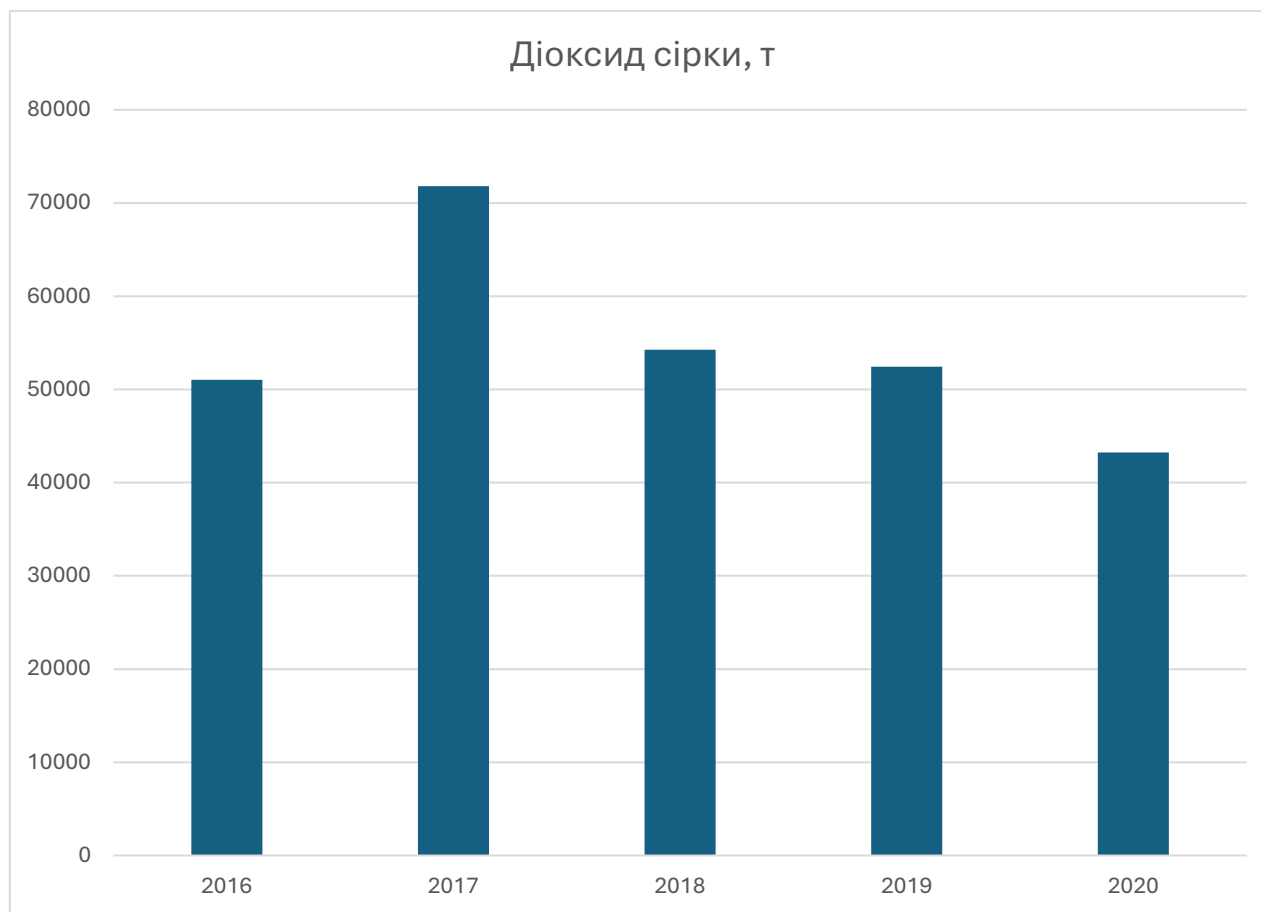


Рисунок 3.1 – Викиди діоксиду сірки від стаціонарних джерел зони „Вінницька“



Рисунок 3.2 – Викиди діоксиду вуглецю від стаціонарних джерел зони „Вінницька“

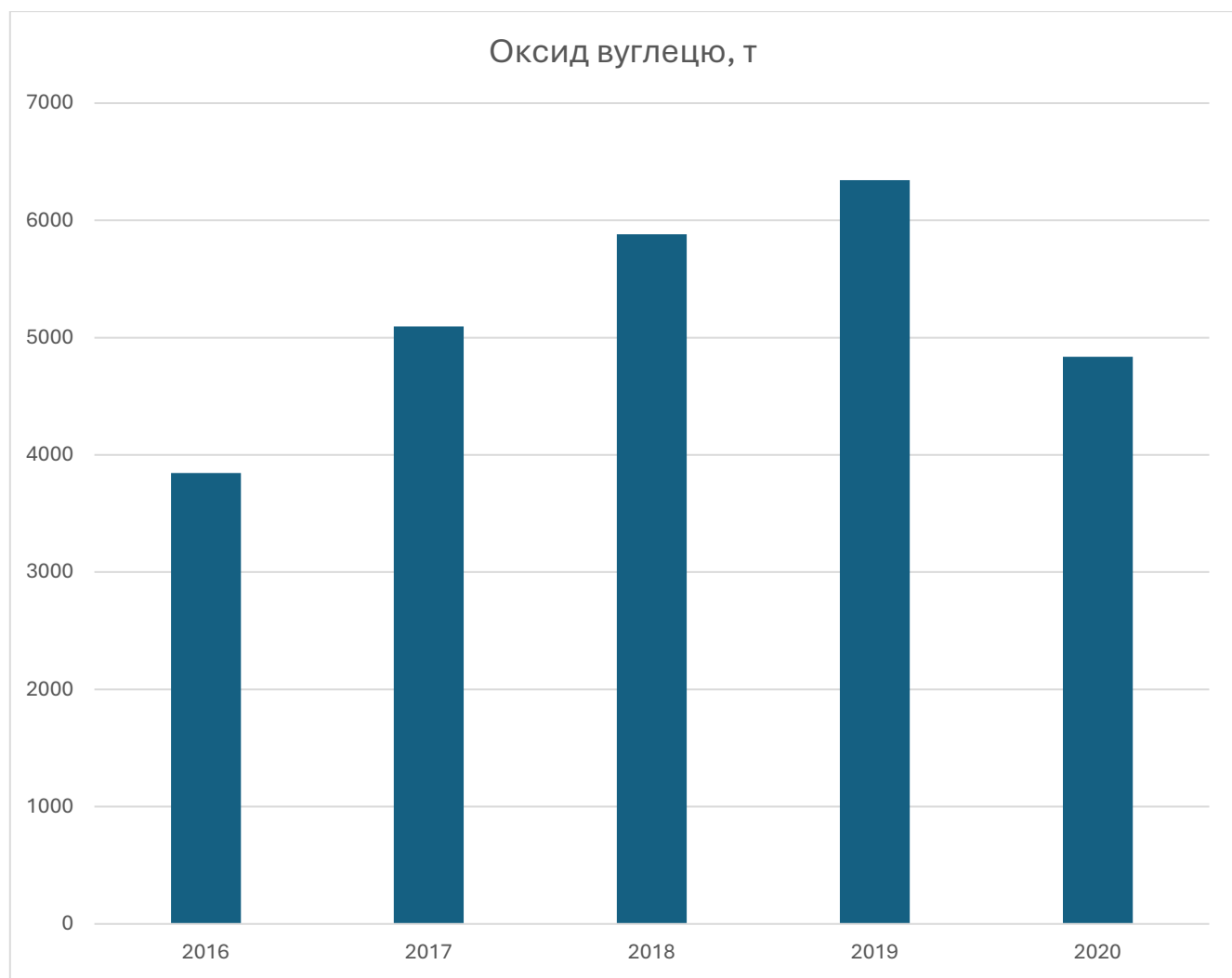


Рисунок 3.3 – Викиди оксиду вуглецю від стаціонарних джерел зони „Вінницька“



Рисунок 3.4 – Викиди речовин у вигляді суспендованих твердих частинок від стаціонарних джерел зони „Вінницька“

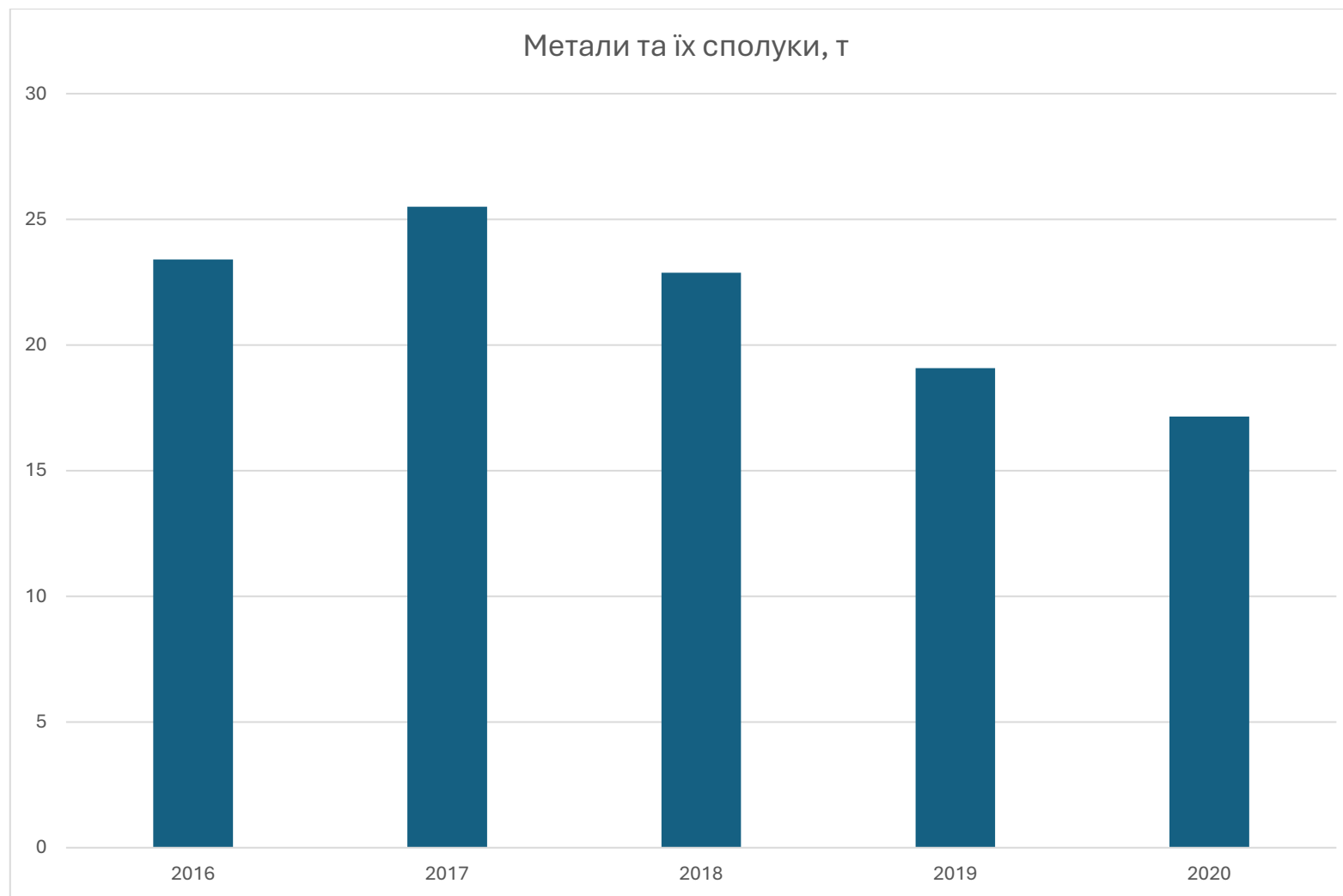


Рисунок 3.5 – Викиди металів та їх сполук від стаціонарних джерел зони „Вінницька“

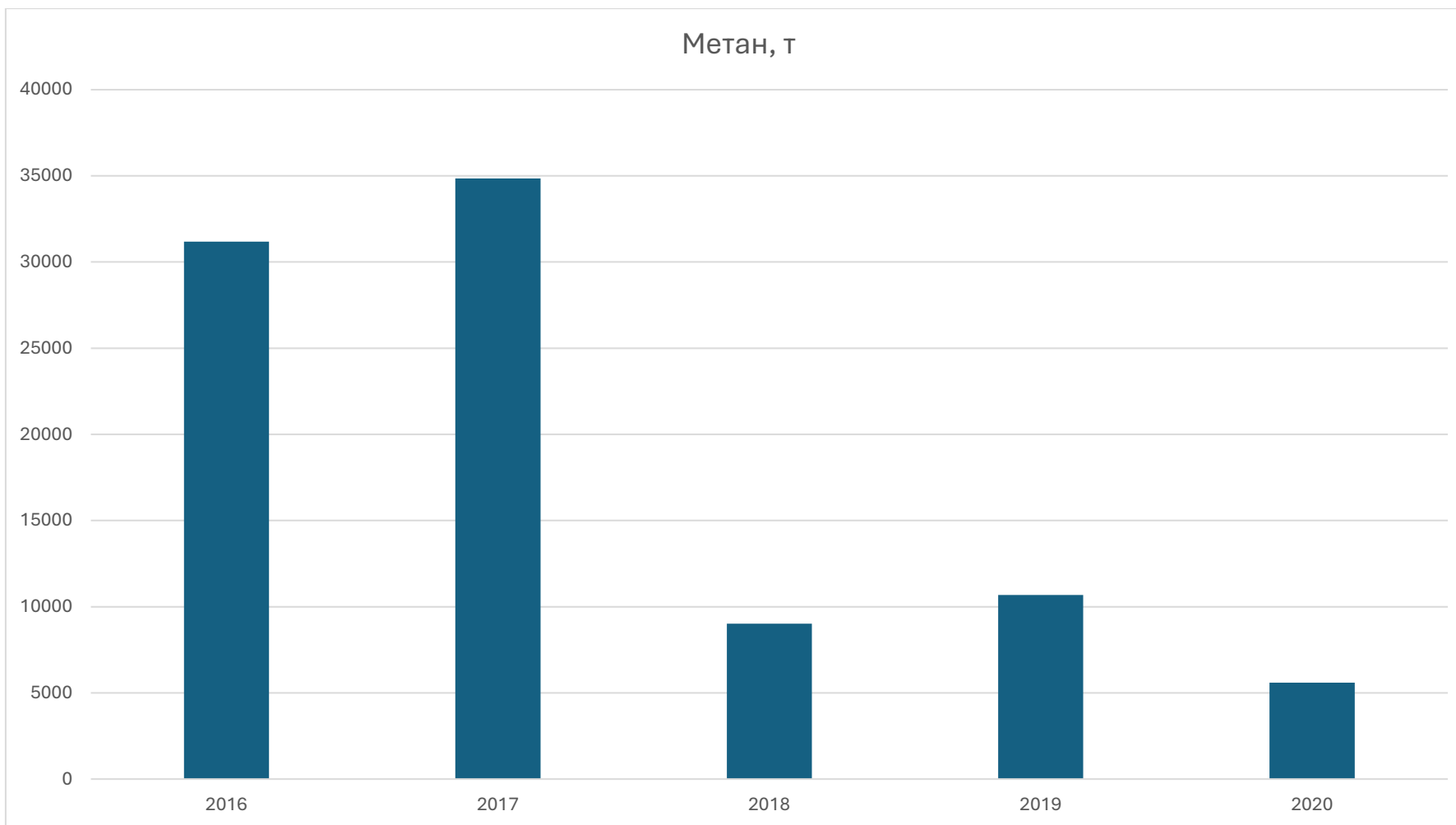


Рисунок 3.6 – Викиди метану від стаціонарних джерел зони „Вінницька“

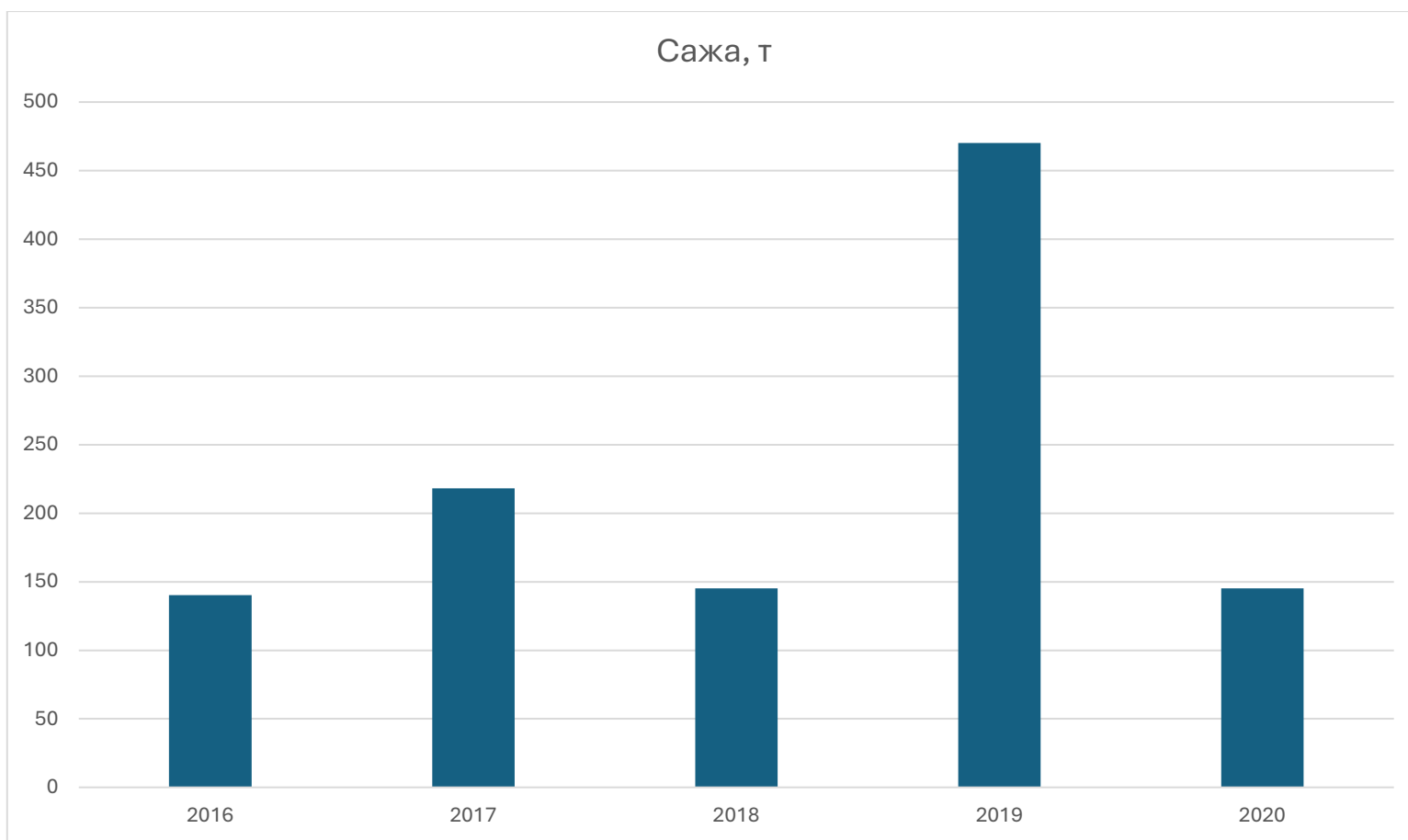


Рисунок 3.7 – Викиди сажі від стаціонарних джерел зони „Вінницька“

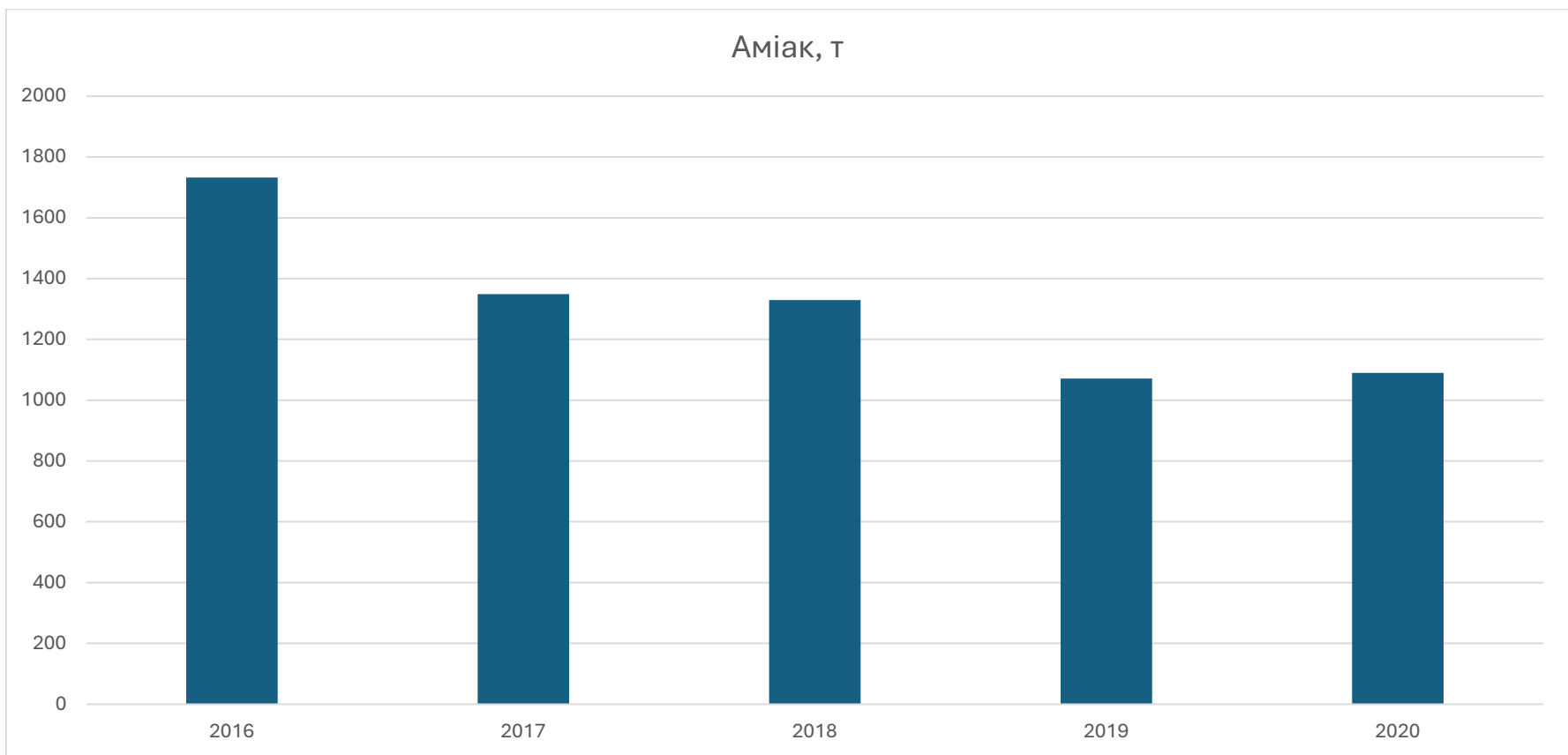


Рисунок 3.8 – Викиди аміаку від стаціонарних джерел зони „Вінницька“

ДОДАТОК І

ЗАПЛАНОВАНІ ЗАХОДИ ЩОДО МОДЕРНІЗАЦІЇ МЕРЕЖІ СПОСТЕРЕЖЕНЬ

Таблиця І.1 - Заплановані заходи щодо встановлення пунктів спостереження за якістю атмосферного повітря

	Етап	Заходи	Строки виконання	Відповідальна установа або організація	Примітки
1	І етап: Впровадження нової державної системи моніторингу у галузі охорони атмосферного повітря	Придбання та встановлення 4 еталонних та 2 індикативних автоматизованих постів спостережень на території зони „Вінницька“.	2022	Управління розвитку територій та інфраструктури Вінницької обласної державної адміністрації, органи місцевого самоврядування (за згодою), територіальні громади (за згодою)	Державний бюджет, обласний бюджет, інші джерела фінансування
	Наукові дослідження щодо необхідності встановлення нових автоматизованих постів спостережень	Проведення досліджень стану атмосферного повітря з метою уточнення режимів оцінювання та оцінювання просторового розподілу концентрацій забруднювальних речовин	2023	Управління розвитку територій та інфраструктури Вінницької обласної державної адміністрації, наукові та інші установи	Державний бюджет, обласний бюджет, інші джерела фінансування
		Оцінювання потреби зони „Вінницька“ у додаткових пунктах спостережень за станом атмосферного повітря та місць їх розташування. Визначення маршрутів для пересувної лабораторії для спостереження за станом атмосферного повітря.		Управління розвитку територій та інфраструктури Вінницької обласної державної адміністрації, наукові та інші установи.	

Продовження таблиці І.1

№	Етап	Заходи	Строки виконання	Відповідальна установа або організація	Примітки
		Формування бази даних діючих дозволів на викиди забруднювальних речовин в атмосферне повітря від стаціонарних джерел забруднення суб'єктів господарювання на території агломерації «Вінниця»	2022-2023	Управління розвитку територій та інфраструктури Вінницької обласної державної адміністрації, наукові та інші установи.	Державний бюджет, обласний бюджет
2	II етап: Розгалуження державної системи моніторингу у галузі охорони атмосферного повітря	Внесення змін до Програми у частині місць встановлення нових постів спостереження за станом атмосферного повітря та їх кошторисної вартості	2023	Управління розвитку територій та інфраструктури Вінницької обласної державної адміністрації	-
		Придбання пересувної лабораторії для спостереження за станом атмосферного повітря.	2023-2025	Управління розвитку територій та інфраструктури Вінницької обласної державної адміністрації, органи місцевого самоврядування (за згодою), територіальні громади (за згодою)	Державний бюджет, обласний бюджет, інші джерела фінансування

Продовження таблиці І.1

3	Обслуговування та забезпечення функціонування системи моніторингу атмосферного повітря зони «Вінницька»	3.1. Обслуговування системи моніторингу атмосферного повітря зони «Вінницька»	2023-2025	Управління розвитку територій та інфраструктури Вінницької обласної державної адміністрації, органи місцевого самоврядування (за згодою), територіальні громади (за згодою)	Державний бюджет, обласний бюджет, інші джерела фінансування
		3.2. Проведення досліджень якості атмосферного повітря тощо			
4	Створення та обслуговування інформаційно-аналітичної системи даних про якість атмосферного повітря зони «Вінницька»	Розроблення, впровадження та забезпечення функціонування інформаційно-аналітичної системи даних про якість атмосферного повітря зони «Вінницька»	2023-2025	Управління розвитку територій та інфраструктури Вінницької обласної державної адміністрації, органи місцевого самоврядування (за згодою), територіальні громади (за згодою)	Державний бюджет, обласний бюджет, інші джерела фінансування

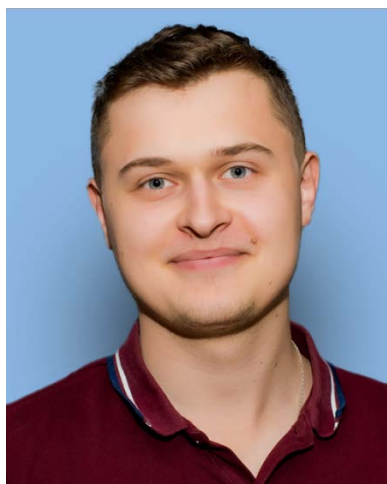
ІНФОРМАЦІЯ ПРО АВТОРІВ



Мокін Віталій Борисович (03 вересня 1974 року) — український вчений в галузі інформаційних технологій і системного аналізу та їх застосування для вирішення задач у різних складних системах, у т.ч. у сфері екології та природних ресурсів. Доктор технічних наук, професор, завідувач кафедри системного аналізу та інформаційних технологій (з 2004 р.), науковий керівник науково-дослідної лабораторії екологічних досліджень та екологічного моніторингу (НДЛ ЕДЕМ) (з 2000 р.), науковий співкерівник НДЛ штучного інтелекту та когнітивістики (з 2023 р.) Вінницького національного технічного університету. У 2009-2010 був радником міністра Мінприроди. Наукові інтереси у сфері охорони довкілля та природокористування – збирання, аналіз, моделювання та прогнозування даних екологічного моніторингу, оптимізація природокористування та управління природними ресурсами, розвиток та застосування геоінформаційних технологій та штучного інтелекту. Брав участь в якості експерта/національного експерта упродовж 2007-2021 рр. в 12 міжнародних проєктах TACIS, EU, UNEP, UNDP, ENVSEC, OSCE, GEF, SIDA, EEF та ін. Науковий керівник більше 50 виконаних науково-дослідних робіт упродовж 2000-2024 рр. Науковий керівник 9 захищених кандидатських дисертацій та дисертацій на здобуття ступеню PhD у сфері інформаційних технологій. Автор біля 500 наукових та навчально-методичних публікацій, зокрема: більше 30 підручників та посібників, 16 монографій. Має більше 50 свідоцтв про реєстрацію авторських прав на твори (комп'ютерні програми). Був науковим керівником робіт з розроблення програм державного моніторингу атмосферного повітря Вінницької агломерації (м. Вінниця) та Вінницької зони (області) (2021 р.).



Крижановський Євгеній Миколайович (16 вересня 1983 року) — український вчений в галузі інформаційних технологій і системного аналізу та їх застосування для вирішення задач у різних складних системах, у т.ч. у сфері екології та аналізу стану атмосферного повітря. Кандидат технічних наук зі спеціальності «Інформаційні технології», магістр зі спеціальності «Екологія та охорона навколишнього природного середовища», доцент, заступник завідувача кафедри системного аналізу та інформаційних технологій Вінницького національного технічного університету. Наукові інтереси у сфері атмосферного повітря— аналіз даних моніторингу стану атмосфери, оптимізація мереж моніторингу атмосфери, розвиток та застосування геоінформаційних технологій, створення геоінформаційних систем моніторингу атмосферного повітря. Брав участь в якості експерта/національного експерта впродовж 2000-2021 рр. в 7 міжнародних проєктах TACIS, EU, UNEP, UNDP, ENVSEC, OSCE, GEF, SIDA та ін. Брав участь у більше 40 науково-дослідних робіт упродовж 2007 – 2024 рр., у т.ч. в одній був науковим керівником, а в більшості – відповідальним виконавцем. Автор більше 170 наукових та навчально-методичних публікацій, у т.ч. біля 20 навчально-методичних. Має більше 30 свідоцтв про реєстрацію авторських прав на твори (комп’ютерні програми). Брав участь у розробленні програм державного моніторингу атмосферного повітря Вінницької агломерації (м. Вінниця) та Вінницької зони (області) (2021 р.)



Шмундяк Дмитро Олександрович (12 травня 1997 року) — український вчений в галузі інформаційних технологій і системного аналізу та їх застосування для вирішення задач у різних складних системах, у т.ч. у сфері забруднення атмосферного повітря. Магістр зі спеціальності «Інформаційні системи та технології», аспірант кафедри системного аналізу та інформаційних технологій Вінницького національного технічного університету зі спеціальності «Системний аналіз». Призер II туру Всеукраїнського конкурсу студентських наукових робіт «Інформатика та кібернетика». Співавтор ряду наукових публікацій, у т.ч. 5 наукових статей у фахових виданнях України та 1 статті в міжнародному виданні, що входить в базу Scopus та Web of Science, за тематикою застосування системного аналізу до систем управління станом атмосферного повітря та інших природних середовищ. Співавтор звіту госпдоговірної науково-дослідної роботи на тему «Розробка розділів програми державного моніторингу у галузі охорони атмосферного повітря Вінницької області», за результатами якої була розроблена Програма державного моніторингу у галузі охорони атмосферного повітря зони «Вінницька» за сучасними вимогами європейських директив та українського законодавства.

Електронне наукове видання

**Шмундяк Дмитро Олександрович,
Мокін Віталій Борисович,
Крижановський Євгеній Миколайович**

**СИСТЕМНИЙ АНАЛІЗ СТАНУ АТМОСФЕРНОГО
ПОВІТРЯ РЕГІОНУ, З УРАХУВАННЯМ ВПЛИВУ
АНОМАЛІЙ**

Монографія

Оригінал-макет підготовлено *Д. Шмундяком*

Видання виготовлено у *РВВ ВНТУ*

Підписано до видання 29.07.2025 р.

Формат 29,7×42¼.

Гарнітура Times New Roman.

Зам № P2025-103.

Видавець та виготовлювач –
Вінницький національний технічний університет,
Редакційно-видавничий відділ.
21021, м. Вінниця, Хмельницьке шосе, 95,
ВНТУ, ГНК, к. 114.
press.vntu.edu.ua
Email: rvv.vntu@gmail.com

Свідоцтво суб'єкта видавничої справи
серія ДК № 3516 від 01.07.2009 р.