

Вінницький національний технічний університет
Факультет комп'ютерних систем і автоматики
Кафедра системного аналізу та інформаційних технологій

МАГІСТЕРСЬКА КВАЛІФІКАЦІЙНА РОБОТА

на тему:

“Інформаційна технологія аналізу та прогнозування якості атмосферного повітря у місті Вінниці за даними громадського моніторингу (комплексна МКР). Частина II. Інтерполювання даних та вибір місць для нових пунктів”

Виконала: студентка 2 курсу, групи 2ІСТ-20м спеціальності 126 – «Інформаційні системи та технології»

_____ Князюк А.О.

Керівник: д.т.н., проф. каф. САІТ

_____ Мокін О.Б.

«__» _____ 2021 р.

Опонент: доц. каф. АІІТ

_____ Іванов Ю.Ю.

«__» _____ 2021 р.

Допущено до захисту

Завідувач кафедри САІТ

_____ д.т.н., проф. Мокін В. Б.

«__» _____ 2021 р.

Вінниця ВНТУ – 2021 рік

Вінницький національний технічний університет
Факультет комп'ютерних систем і автоматики
Кафедра системного аналізу та інформаційних технологій
Рівень вищої освіти – II-й (магістерський)
Галузь знань – 12 Інформаційні технології
Спеціальність – 126 Інформаційні системи та технології
Освітньо-професійна програма – Інформаційні технології аналізу даних та зображень

ЗАТВЕРДЖУЮ

Завідувач кафедри САІТ

_____ д.т.н., проф. Мокін В. Б.

«__» _____ 2021 р.

ЗАВДАННЯ
НА МАГІСТЕРСЬКУ КВАЛІФІКАЦІЙНУ РОБОТУ СТУДЕНТЦІ
Князюк Анастасії Олексівні

1. Тема роботи: “Інформаційна технологія аналізу та прогнозування якості атмосферного повітря у місті Вінниці за даними громадського моніторингу (комплексна МКР). Частина II. Інтерполювання даних та вибір місць для нових пунктів”,

керівник роботи: Мокін О.Б., д.т.н., проф. каф. САІТ,

затверджені наказом закладу вищої освіти від «__» _____ 2021 року №__

2. Строк подання студенткою роботи «__» _____ 2021 року

3. Вихідні дані до роботи:

Набір даних моніторингу атмосферного повітря по 9 постах наданими ресурсами Eco City та SaveEcoBot.

4. Зміст текстової частини:

- аналіз сучасного стану моніторингу атмосферного повітря;
- основні етапи виконання роботи та огляд набору вхідних даних;
- обробка результатів спостережень та візуалізація даних у середовищі Kaggle;
- економічна частина.

5. Перелік ілюстративного матеріалу (з точним зазначенням обов'язкових креслень):

- сферична варіограма по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 26.11.2021 о 00:00:00;
- сферична варіограма по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 26.11.2021 о 01:00:00;
- сферична варіограма по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 26.11.2021 о 02:00:00;

- сферична варіограма по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 26.11.2021 о 03:00:00;
- сферична варіограма по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 26.11.2021 о 04:00:00;
- сферична варіограма по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 26.11.2021 о 05:00:00;
- сферична варіограма по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 26.11.2021 о 06:00:00.

6. Консультанти розділів МКР

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
3	Мокін В.Б., д.т.н., проф. каф. САІТ		
4	Ратушняк О.Г., к.т.н., доц. каф. ЕПВМ		

7. Дата видачі завдання « ___ » _____ 2021 року

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів МКР	Строк виконання етапів роботи	Примітка
1	Аналіз сучасного стану моніторингу атмосферного повітря	09.2021	
2	Основні етапи виконання роботи та огляд набору вхідних даних	09.2021	
3	Обробка результатів спостережень та візуалізація даних у середовищі Kaggle	10.2021	
4	Економічна частина	10.2021	
7	Оформлення матеріалів до захисту МКР	11.2021	

Студентка _____

Князюк А О.

Керівник роботи _____

Мокін О. Б.

АНОТАЦІЯ

УДК 004.08

Князюк А. О. Інформаційна технологія аналізу та прогнозування якості атмосферного повітря у місті Вінниці за даними громадського моніторингу (комплексна МКР). Частина II. Інтерполювання даних та вибір місць для нових пунктів. Магістерська кваліфікаційна робота зі спеціальності 126 – інформаційні системи та технології, освітньо-професійна програма – інформаційні технології аналізу даних та зображень. Вінниця: ВНТУ, 2021. 107 с.

На укр. мові. Бібліогр.: 20 назв; рис.: 56; табл.: 7.

В магістерській кваліфікаційній роботі проаналізовано можливість побудови інтерполяційних картограм стану забруднення атмосферного повітря у м. Вінниці за даними громадського моніторингу на Python та запропоновано нову методику вибору місць для нових пунктів спостереження для підвищення ефективності їх геостатистичного аналізу. Об'єктом досліджень є процес моніторингу якості атмосферного повітря у місті Вінниця.

Ілюстративна частина складається з 7 плакатів із результатами геостатистичного аналізу.

У розділі економічної частини розглянуто питання про доцільність розробки та впровадження інформаційної технології визначення місць для нових постів моніторингу атмосферного повітря.

Ключові слова: моніторинг якості атмосферного повітря міста, інтерполяція, геостатистичний аналіз, громадський моніторинг, аналіз та прогнозування даних.

ABSTRACT

Kniazuk A. A. Information technology of analysis and forecasting of atmospheric air quality in the city of Vinnytsia according to public monitoring (complex MCR). Part II. Data interpolation and selection of locations for new items. Master's thesis in specialty 126 - information systems and technologies, educational and professional program - information technology data and image analysis. Vinnytsia: VNTU, 2021. 107 p.

In Ukrainian language. Bibliogr .: 20 titles; fig .: 56; table: 7.

The master's qualification work analyzes the possibility of constructing interpolation maps of air pollution in Vinnytsia according to public monitoring on Python and proposes a new method of site selection for new observation points to increase the efficiency of their geostatistical analysis. The object of research is the process of monitoring air quality in the city of Vinnytsia.

The illustrative part consists of 7 posters with the results of geostatistical analysis.

In the section of the economic part the question of expediency of development and introduction of information technology of definition of places for new posts of monitoring of atmospheric air is considered.

Key words: city air quality monitoring, interpolation, geostatistical analysis, public monitoring, data analysis and forecasting.

ЗМІСТ

ВСТУП	4
1 АНАЛІЗ СУЧАСНОГО СТАНУ МОНІТОРИНГУ АТМОСФЕРНОГО ПОВІТРЯ.....	6
1.1 Аналіз предметної області.....	6
1.2 Пости спостереження.....	8
1.3 Сучасний стан моніторингу у місті Вінниця.....	16
1.4 Особливості прогнозування рівня забрудненості.....	20
1.5 Висновки	21
2 ОСНОВНІ ЕТАПИ ВИКОНАННЯ РОБОТИ ТА ОГЛЯД НАБОРУ ВХІДНИХ ДАНИХ	22
2.1 Вибір оптимальних інформаційних технологій.....	22
2.1.1 ArcGIS Online.....	25
2.1.2 Kriging інтерполяція.....	27
2.1.3 Kaggle	29
2.1.4 Приклади застосування інтерполяції	32
2.1.5 Інструменти інтерполяції	34
2.1.6 Порівняння методів інтерполяції.....	36
2.2 Огляд вхідного набору даних	41
2.3 Висновки	53
3 ОБРОБКИ РЕЗУЛЬТАТІВ СПОСТЕРЕЖЕНЬ ТА ВІЗУАЛІЗАЦІЇ ДАНИХ У СЕРЕДОВИЩІ KAGGLE	54
3.1 Підготовка вхідних даних	54
3.2 Геостатистичний аналіз за допомогою SciKit-GStat	57
3.3 Лінійна інтерполяція за допомогою scipy.interpolate.interp2d.....	70
3.4 Висновки	71
4 ЕКОНОМІЧНА ЧАСТИНА	72
4.1 Оцінювання комерційного потенціалу розробки.....	72
4.2 Прогнозування витрат на виконання науково-дослідної роботи	75
4.3 Розрахунок економічної ефективності науково-технічної розробки	80

4.4 Розрахунок ефективності вкладених інвестицій та періоду їх окупності .	82
4.5 Висновки	85
ВИСНОВКИ.....	86
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	88
Додаток А (обов'язковий). Технічне завдання	92
Додаток Б (обов'язковий). Протокол перевірки кваліфікаційної роботи	95
Додаток В (довідниковий). Лістинг програми	96
Додаток Г (обов'язковий). Ілюстративна частина	100

ВСТУП

Актуальність теми. Забруднення атмосферного повітря є одним з важливих факторів, що впливають на здоров'я людини. Через високий рівень забрудненості повітря більшість населення страждають захворюванням серця, хвороби пов'язані з інсультом, раком легень, хронічними та гострими респіраторними захворюваннями, включаючи астму. За рахунок заходів щодо зниження рівня забруднення повітря країни можуть зменшити тягар хвороби. Однак забруднення повітря не отримує належної уваги ні на національному, ні міжнародному рівні. Наразі у місті Вінниця існують пости вимірювання якості атмосферного повітря, проте є ділянки міста в яких це не контролюється, дізнатись про рівень забрудненості там складніше та результати будуть з меншою точністю.

Мета і завдання роботи. Метою даної роботи є поліпшення точності побудови картограм реального чи прогнозованого стану забруднення атмосферного повітря у м. Вінниці за даними громадського моніторингу шляхом розроблення методики вибору місць для нових пунктів спостереження. За допомогою спеціального середовища, можливо здійснити аналіз даних та визначення місць для нових постів, аби покращити моніторинг рівня забрудненості атмосферного повітря.

Розробка системи визначення нових місць передбачає виконання задач таких етапів:

- розвідувальний аналіз даних, отриманих з постів розміщених у місті Вінниця за певний період;
- вибір середовища та оптимальних технологій;
- підготовка набору даних для розміщення у спеціальному середовищі на карті;
- лінійна інтерполяція даних.

Об'єктом дослідження магістерської кваліфікаційної роботи є процес моніторингу якості атмосферного повітря у місті Вінниця.

Предметом дослідження магістерської кваліфікаційної роботи є технології та засоби визначення нових місць для постів моніторингу забрудненості атмосферного повітря.

Методи дослідження. У дослідженнях використовувались методи ГІС-технологій для формалізації просторових даних ГІС, методи реляційних баз даних для формалізації та збереження атрибутивних даних ГІС. Також було здійснено аналіз даних та інтерполяцію даних у середовищі Kaggle, використовуючи мову програмування Python.

Новизна одержаних результатів. Дістала подальший розвиток інформаційна технологія інтерполювання якості атмосферного повітря за методом геостатистичного аналізу з використанням ArcGIS та Python-бібліотек, що дозволило дослідити ряд сценаріїв і визначити можливі місця розташування нових постів, що дозволить покращити точність цього інтерполювання на всю територію міста Вінниці.

Практичне значення роботи сприятиме покращенню моніторингу атмосферного повітря у місті. Результати роботи мають цінність для екологічного контролю забруднення атмосферного повітря.

Апробація результатів магістерської кваліфікаційної роботи. Результати роботи доповідались на всеукраїнській науково-практичній інтернет-конференції «Молодь в науці: дослідження, проблеми, перспективи» (Вінниця, 2021-2022 рр.).

Публікації результатів магістерської кваліфікаційної роботи. Опубліковано тези на Всеукраїнській науково-практичній інтернет-конференції «Молодь в науці: дослідження, проблеми, перспективи» (Вінниця, 2021-2022 рр.) [1].

1 АНАЛІЗ СУЧАСНОГО СТАНУ МОНІТОРИНГУ АТМОСФЕРНОГО ПОВІТРЯ

1.1 Аналіз предметної області

Атмосферне повітря є важливим компонентом природнього довкілля, та від якості повітря залежать умови життя на Землі. В наш час стан атмосферного повітря постійно погіршується різними забруднюючими факторами і викидами та впливає на рівень життя населення. Саме людська діяльність призводить до змін атмосфери. Для того, аби зменшити шкідливий вплив на атмосферу повітря потрібно здійснювати постійний моніторинг його стану.

Якість повітря залежить від ступеня забруднення повітря, тобто концентрації забруднюючих речовин, що містяться в ньому. Якість повітря не тільки змінюється щодня, але й змінюється протягом кількох годин. Це пояснюється тим, що ступінь забруднення змінюється залежно від кількості викидів в атмосферу. Наприклад, починаючи з часу доби, забруднення повітря є найвищим у години пік. З огляду на погодні умови та сильний вітер, якість повітря зазвичай є значною. краще. У зв'язку з цим необхідно вимірювати ступінь забруднення повітря [6].

Незважаючи на існування великої кількості різних забруднюючих речовин, 98% викидів складаються лише з 4 забруднюючих речовин: твердих частинок, вуглекислого газу, оксидів азоту та оксиду вуглецю. Крім того, концентрації цих речовин часто перевищують допустимі норми. Дані про концентрацію різних речовин додаються разом, щоб отримати інформацію про ступінь забруднення повітря. Ступінь небезпеки для людини такого рівня забруднення визначається згідно з нормами, встановленими національним наглядовим органом.

Для обчислення індексу забруднення необхідна інформація про концентрацію забруднюючих речовин у атмосфері за певний період часу. Цю

інформацію зазвичай отримують з моніторингу повітря. Для перетворення концентрації в індекс забруднення використовується спеціальна математична функція, коефіцієнт якої залежить від типу забруднюючої речовини. Це тому, що різні речовини можуть бути небезпечними в різному ступені. Отримане значення потім оцінюється за певною шкалою, де значення показника зазвичай групують у певний діапазон. Кожен діапазон має назву, опис можливих наслідків для здоров'я та колірний код у деяких індексах. На основі значення індексу якості повітря органи влади можуть робити різні пропозиції для населення щодо зменшення шкідливого впливу забруднюючих речовин на організм [2].

Моніторинг атмосферного повітря – це система спостережень за станом атмосфери, його забрудненням і природними явищами, які відбуваються в ньому, а також оцінка і прогноз стану атмосферного повітря (контроль, аналіз, висновки). У багатьох містах в розвинених країнах світу вже використовують такі пости, протягом останній десятиліть.

Моніторинг у галузі охорони атмосферного повітря проводиться з метою отримання, збирання, оброблення, збереження та аналізу інформації про рівень забруднення атмосферного повітря, оцінки та прогнозування його змін і ступеня небезпечності та розроблення науково обґрунтованих рекомендацій для прийняття рішень у галузі охорони атмосферного повітря.

Моніторинг атмосфери включає такі етапи:

- визначення мети та об'єктів спостереження;
- спостереження за джерелами забруднення;
- встановлення основних видів забруднення і забруднюючих речовин;
- вивчення впливу забруднення атмосфери на живі організми;
- прогноз змін середовища в результаті забруднення атмосферного повітря;
- розробка заходів і прийняття управлінських рішень, які направлені на збереження повітря.

Завдання моніторингу довкілля вимагають участі сучасних обчислювальних засобів, які потребують обробки великого обсягу інформації, особливо її класифікації, попередньої обробки та подальшого використання в складних обчислювальних алгоритмах для дослідницьких процесів, пов'язаних з перенесенням забруднюючих речовин. Використання даних спостережень та модельних уявлень процесу поширення домішок для створення моделей для оцінки полів концентрації та параметрів джерела дозволяє більш надійно контролювати основні параметри техногенного забруднення. Цей метод дозволяє визначити обсяг інформації в системі спостереження та оптимізувати розташування та кількість точок відбору проб. Завдання виявлення джерела забруднення є актуальним, оскільки його вирішення дозволяє визначити внесок окремого джерела (підприємства, міста, країни) у забруднення повітря чи ґрунту на даний момент. Це необхідно враховувати при створенні нових джерел забруднення – запровадженні нових промислових підприємств та управлінні викидами діючих підприємств, щоб їх загальна кількість не перевищувала встановлених нормативів.

1.2 Пости спостереження

Постом спостереження можна назвати точку місцевості (вибране місце), на якому розміщують прилад, що фіксує рівень атмосферного повітря з певною періодичністю та передає дані до лабораторії [3]. Пости спостережень за забрудненнями можуть бути стаціонарними, маршрутними та пересувними (підфакельними). На рисунку 1.1 зображений стаціонарний пост спостереження.

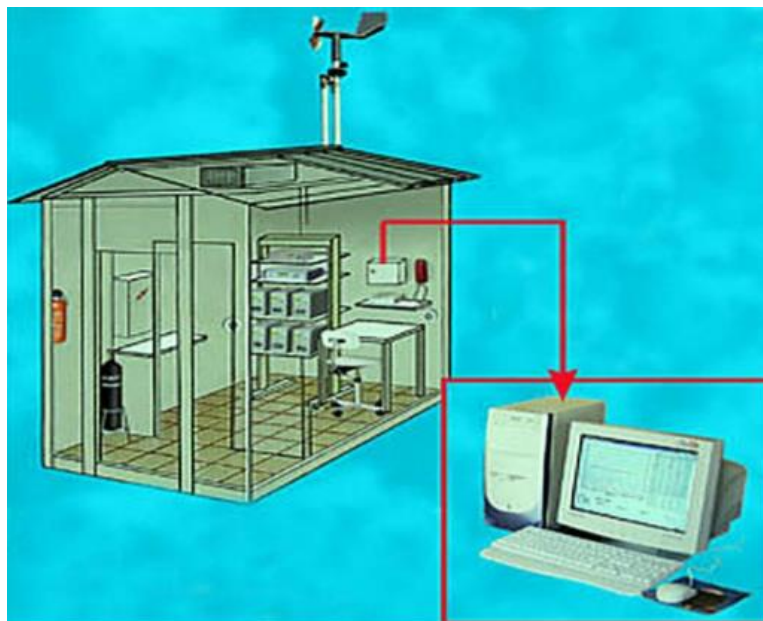


Рисунок 1.1 – Стаціонарний пост спостереження АТМОСФЕРА-10

Лабораторію використовують для стаціонарних спостережень за рівнем забруднення атмосферного повітря, а також для з'ясування метеорологічних характеристик [4].

На рисунку 1.2 зображена пересувна екологічна лабораторія ПЕЛКЗА, або ж маршрутний пост спостереження.



Рисунок 1.2 – Пересувна екологічна лабораторія ПЕЛКЗА

Маршрутний пост спостереження призначений для регулярного відбору проб повітря у фіксованих точках місцевості за допомогою спеціально обладнаної автолабораторії. Маршрут щомісячно змінюється з таким розрахунком, щоб відбір проб повітря у кожному пункті проводився в різний час доби. Розміщення маршрутних постів повинно бути таким, щоб виявляти максимальні концентрації забруднюючих речовин, які формуються джерелом викиду. Визначаючи місця відбору проб, приймають до уваги висоту джерела викиду і максимально можливу зону забруднення ним атмосферного повітря.

Пересувний (підфакельний) пост (пересувна комплектна лабораторія) призначений для відбору проб під димовим (газовим) факелом з метою виявлення зони впливу даного джерела промислових викидів. Підфакельні пости є точками, розташованими на фіксованих відстанях від джерела. Вони переміщуються відповідно до напрямку факела джерела викидів.

При виборі місця розміщення поста потрібно встановити, яку інформацію необхідно отримати:

- а) рівень забруднення атмосфери, характерний для даного району міста;
- б) концентрацію домішок в конкретній точці, що перебуває під впливом викидів окремого промислового підприємства, великої автомагістралі [5].

Перед встановленням посту слід досконало вивчити: розрахункові поля концентрацій по всіх компонентах від сукупності викидів всіх стаціонарних і пересувних джерел; особливості забудови та рельєфу території: можливості становлення житлової забудови та розширення підприємств промисловості, енергетики, комунального господарства; автотранспорту та інших секторів економіки міського господарства, активні особливості обраної зони; щільність населення; метеорологічні обставини даної території та інших. Пост повинен перебувати поза аеродинамічної тіні будинків та зони зелених насаджень, його земля повинна добре провітрюватися, не піддаватися впливу прилеглих низьких джерел (стоянок автомобілів, невеликих підприємств з низькими викидами тощо.). Чисельність стаціонарних постів у будь-якому місті (населеному пункті) визначається кількістю населення, рельєфом території,

особливостями промисловості, активною структурою (житлова, промислова, зелена зона тощо.), просторової і тимчасової мінливості полів концентрацій шкідливих препаратів. Приміром, з кількості населення, чисельність постів орієнтується наступним чином (табл. 1.1).

Таблиця 1.1 – Залежність кількості стаціонарних постів від чисельності населення

Чисельність населення, тис. чол.	Кількість постів
<50	1
50-100	2
100-200	3
200-500	3-5
500-1000	5-10
1000-2000	10-15
>2000	15-20

Для населених пунктів зі складним рельєфом і великою кількістю джерел рекомендовано ставити раз на будь-які (5-10) км². Щоб інформація про забруднення повітря враховувала особливості мегаполісу, рекомендовано призначати посади досліджень у різноманітних активних зонах – житлової, промислової та зони розваг. У містах з величезною інтенсивністю руху транспорту пости уточнюються ще й близько автомагістралей.

Для забезпечення відповідних умов проведення стаціонарних досліджень українською промисловістю випускаються нормальні павільйони-пости досліджень або комплектні лабораторії типу ПОСТ. Лабораторія ПОСТ – це утеплений, оббитий алюмінієвими осередками павільйон, в якому поставлені комплекти пристроїв та обладнання для відбору проб повітря, проведення метеорологічних вимірювань: швидкості та спрямованості вітру, температури, вологи [6]. Буквально усі стаціонарні пункти контролю забруднення обладнані комплектними лабораторіями ПОСТ-1. Випускаються і уточнюються більш нові трансформації модифікації лабораторії – ПОСТ-2 і

ПОСТ-2а, які виділяються більш високою продуктивністю відбору проб і ступенем автоматизації.

На стаціонарних постах дослідження за забрудненням атмосферного повітря та метеорологічними параметрами повинні проводитися щорічно, у всі сезони, автономно від погодних умов. Для постів досліджень, як правило, уточнюються три програми дослідження: абсолютна, неповна і скорочена. За абсолютною програмою дослідження ведуться щодня (вихідні-неділі, суботи - чергуються) в 1, 7, 13 і 19 годин районного декретного часу, або за ковзним графіком: вівторок, четвер, субота – 7, 10 і 13 год; Перший день тижня, середа, п'ятниця – 15, 18 і 21 год. Дослідження по 1 програмі передбачають вимірювання вмісту в повітрі як провідних, так і своєрідних забруднюючих препаратів. По неповній програмі дослідження ведуться щодня (неділі та суботи чергуються), але тільки в 7, 13 і 19 год районного декретного часу [7].

У районах, де температура повітря нижче 45°C, дослідження ведуться за скороченою програмою щодня, крім неділі, в 7 і 13 год за місцевим декретним часом. Дослідження за скороченою програмою допускається проводити ще в просторах, де середні місячні зосередженні менше 1/20 ПДК_{мр} або менше нижньої межі діапазону вимірювань домішки використовуваним методом.

При не дуже сприятливих метеорологічних умовах (туман, тривала інверсія температур та ін.) відбір проб повітря на всіх постах досліджень повинен виконуватися через будь-які 3 год. Одночасно йде по стопах віднімати проби під смолоскипами провідних джерел забруднення на землі більшої щільності населення. Підфакельні дослідження виконуються за відповідними для наданого фірми домішками.

Стаціонарне місце контролю радіоактивного забруднення атмосферного повітря являє собою або стаціонарний павільйон на кшталт ПОСТ, або будиночок об'ємом 3х3х3 м. Він уточняється, як правило, на спеціально оснащених гідрометеорологічних станціях (ГМС), обгороджених залізною мережею з площею 1 5х10 м, а висота сітки – 1,2-1,5 м. Майданчик повинен перебувати на відстані не менше 10 висот до близької будівлі і не менше 30 м

від доріг. Майданчик повинен мати трав'яний покрив. Не допускається висаджування інших рослин, що більше чагарників та дерев.

На землі ГМС не ближче 4 м від будиночка та огорожі уточняється марлевий планшет для збирання радіоактивних випадіннь та термомлюмінесцентний дозиметр. Установку для відбору проб повітря ніж будь-якого іншого розташовувати в спеціальній будці з жалюзі, піднятою над поверхнею ґрунту на 80 – 100 см. Викид повітря, що пройшов через фільтри установки типу "Тайфун", повинен виконуватися в обов'язковому порядку в зворотну від планшета сторону. Якщо стаціонарне місце не забезпечене електроживленням (трифазне (5-10) кВт), то замість фільтруючої установки допускається використання марлевого конуса [8].

Нагляд за радіоактивністю атмосферного повітря виконується постійний цілий рік. Зміна марлі на планшетах і вертикальних екранах, а також фільтрів в установках виконується щодня о 7 год 30 хв ранку за місцевим декретним часом. З фільтруючих установок фільтри можуть фотографуватися як через 24 год – в 7 год 30 хв, – наприклад і через 12 год, тобто. 2 рази на день. При дворазовому відборі встановлено час роботи установок: з 7 год 30 хв до 13 год 30 хв та з 19 год 30 хв до 1 год 30 хв. Швидкість повітря в апараті орієнтується за допомогою витратомірів УС-125 або УС-175-12 3 рази на день: в 7ч 30 хв, 13 год 30 хв і 1 год 30 хв.

Маршрутний пост досліджень – простір на конкретному маршруті у мегаполісі. Він призначений для постійного відбору проб повітря у фіксованій точці території при наглядах, що ведуться за допомогою пересувної апаратури. Маршрутні дослідження виконуються на маршрутних постах із підтримкою автолабораторій. Ця пересувна лабораторія містить продуктивність близько 5000 відборів проб на рік, в день на такій машині можна зробити вибір 8 – 10 проб повітря. Порядок об'їзду маршрутних постів щомісяця змінюється в такий спосіб, щоб відбір проб повітря будь-якому пункті вівся у час дня і ночі. Наприклад, в перший місяць автомат об'їжджає пости в порядку зростання

номерів, в другому – в порядку їх зменшення, а в третій – з середини маршруту до кінця і від початку до середини і т.д.

Пересувний (підфакельний) пост спеціалізований для відбору проб під димовим (газовим) факелом з метою виявлення зони впливу наданого джерела. Підфакельні дослідження здійснюються по спеціально розроблених програмах і маршрутах за специфічними забруднювальними препаратами, характерними для викидів даного підприємства. Простір відбору проб при підфакельних наглядах обирають на різних відстанях від джерела забруднення з урахуванням закономірностей поширення забруднюючих препаратів в атмосфері. Відбір проб повітря виконується по черзі за спрямованістю вітру на відстанях (0,2 – 0,5); 1, 2, 3, 4, 6, 8, 10, 15 і 20 км від стаціонарного джерела викиду, а також з вітряного боку джерела. Дослідження під факелом ведуться за звичайними для наданого фірми компонентами з урахуванням обсягу викидів та його токсичності [9]. У зоні граничного забруднення (за даними розрахунків і дослідних вимірів) відбирається щонайменше 60 проб повітря, а інших зонах мінімальна кількість може бути щонайменше 25. Відбір проб повітря під час проведення підфакельних досліджень виконується на височини 1,5 м від площині землі на напрямок. 20 – 30, хв не менше ніж у трьох точках одночасно. У напрямку робочого дня під факелом можна відібрати проби по черзі в 5 – 8 точках.

Автоматична система досліджень та контролю навколишнього середовища (АНКОС-АГ) призначена для автоматичного збору, обробки та передачі інформації про рівень забруднення атмосферного повітря. Система дозволяє безперервно отримувати інформацію про зосередженні домішок і метеорологічні параметри в населених пунктах або в межах великих промислових підприємств. Технічні можливості реєстрації, передачі, заощадження та обробки даних про забруднення атмосферного повітря дозволили створити основні засади функціонування автоматичних систем дослідження станом атмосферного повітря.

До складу розробленої індустрією АНКОС-АГ входять відповідні технічні методи:

- павільйон, що конструктивно являє собою залізний каркас прямокутної форми об'ємом 2300x4700x7600 мм;
- щоглові прилади з комплектом метеодатчиків, поставлених на даху павільйону, для вимірювання швидкості та спрямованості вітру, температури, вологості;
- приладу опалення, вентиляції, освітлення, кондиціонування та пожежогасіння;
- газоаналізатори оксиду вуглецю, діоксиду сірки, оксиду, діоксиду та суми оксидів азоту, озону, суми вуглеводнів без метану;
- прилад збору та обробки інформації на основі мікроЕОМ [10].

Замін інформацією між системою АНКОС та Центром обробки інформації здійснюється по комутованих телефонних каналах спільного використання за підтримки апаратів передачі даних (АПД) і мультиплексора передачі даних (МПД). АПД, встановлювані на станціях АНКОС, поруч із АПД і МПД Центру обробки інформації утворюють автоматичну централізовану підсистему збору інформації від систем АНКОС, розміщених по місту або ж району. Склад технічних засобів центру обробки інформації:

- спец обчислювальний ансамбль на основі ЕОМ;
- мультиплексор передачі на основі мікроЕОМ;
- пульт диспетчера;
- мнемосхема;
- додаткове та сервісне обладнання;
- програмне забезпечення (пакета програм первинної та вторинної обробки даних вимірювань, банки даних, диспетчерські програми та інших).

Системи АНКОС-АГ та Центру забезпечують:

- постійний вимір даних параметрів атмосферного повітря;
- автоматичний збір інформації зі станцій АНКОС;

- збір інформації від неавтоматизованих ланок досліджень (наприклад, від стаціонарних та пересувних постів);
- оперативну оцінку історії за знайомими значеннями ГДК;
- короткочасний моніторинг значень забруднення контрольованих домішок;
- обробку та видачу інформації.

Методи математичного забезпечення включають відповідні основні способи обробки даних:

- спосіб первинної обробки (перевірка достовірності казенної інформації про забруднення, приведення інформації до вигляду, зручному для обробки та інших.);
- метод статистичної обробки (визначення числових, імовірнісних даних характеристик забруднення, метеорологічних характеристик та ін.);
- метод експрес-інформації про стан забруднення у всіх районах мегаполісу на даний момент часу;
- метод короткочасного та довготривалого прогнозування забруднення повітря;
- метод управління, визначальний тимчасової режим роботи системи, черговість рубежів функціонування, контроль працездатності системи, цінність програм обробки даних та інших.

Час усереднення даних про концентрації домішок становить не менше 20 – 30 хв, що відповідає часу відбору проб в поглинальні прилади. Частота видачі інформації автоматичної системи може оформлювати від декількох хв. до кількох годин [11].

1.3 Сучасний стан моніторингу у місті Вінниця

Основні проблеми з моніторингом якості атмосферного повітря України пов'язані з декількома факторами:

- відсутністю систематичного фінансування;

- відсутністю стратегії розвитку моніторингових спроможностей;
- відсутністю комунікаційної стратегії.

Натомість різні громадські організації створюють проекти до яких можливо долучитись або ж зробити свій внесок у проект [12].

Наприклад, у місті Вінниця є такі проекти, як SaveDnipro, Eco City, Український гідрометеорологічний центр та SaveEcoBot. На рисунку 1.3 можемо бачити сайт проекту SaveEcoBot.

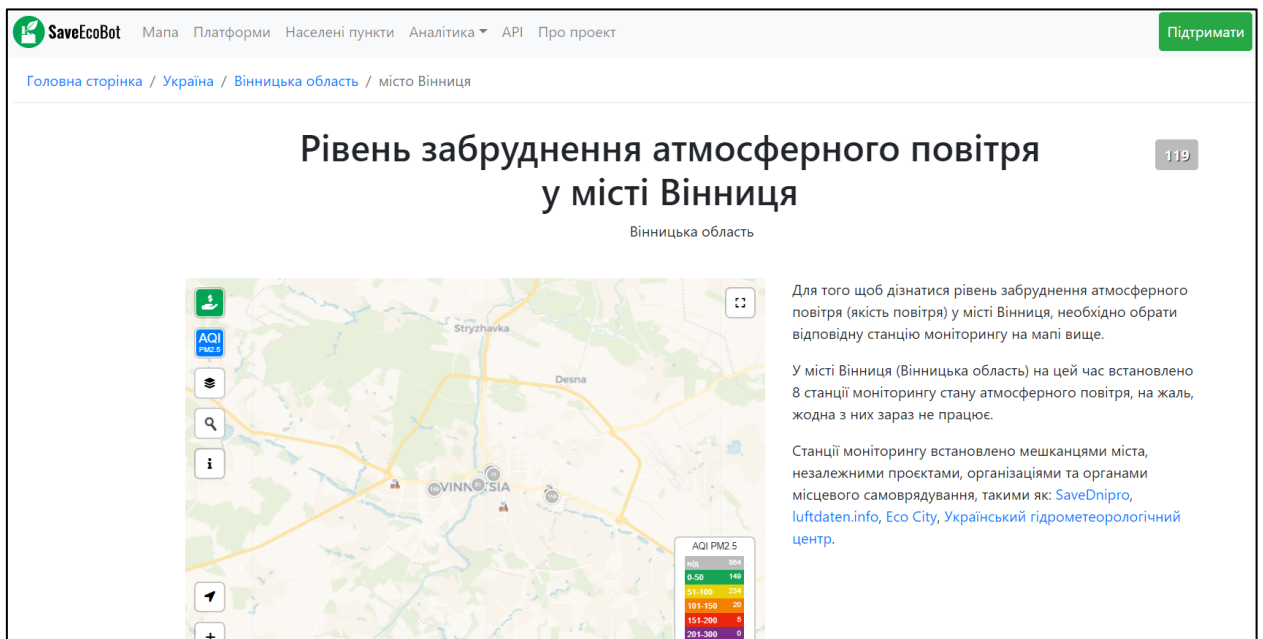


Рисунок 1.3 – Сайт проекту моніторингу якості повітря SaveEcoBot

Станції даного проекту було становлено в таких містах, як Вінниця, Хмільник та Ладижин. Станції встановлюють мешканці міста та органи місцевого самоврядування.

Також можна відзначити проект Eco City, на рисунку 1.4 зображена головна сторінка сайту, де власне відображається одразу ж карта моніторингу якості повітря з постами по місту та області та показники вимірювання за певними ознаками.

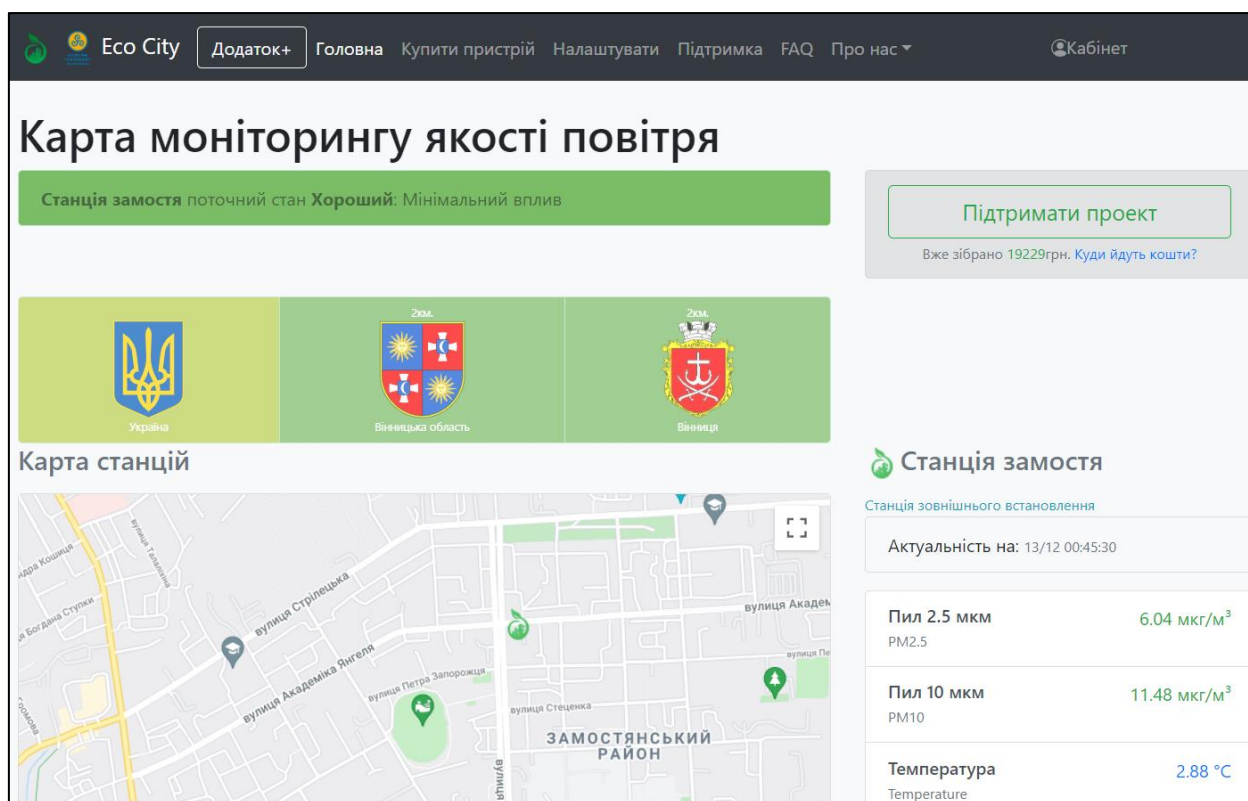


Рисунок 1.4 – Головна сторінка сайту проекту Eco City

Громадський моніторинг якості повітря України EcoCity є соціальним, екологічним, незалежним проектом, який реалізовується неприбутковою громадською організацією "Фрі Ардуіно".

Дані постійно оновлюються та розміщені на головній сторінці того чи іншого проекту, також дані є у вільному доступі, тобто їх можна завантажити або ж оновлювати на сайті.

Наразі у Вінниці розміщено 9 постів спостереження моніторингу якості атмосферного повітря. Більшість постів розміщені у центрі, та є частини міста де їх немає, тож головним завданням даної роботи є визначення місця, де краще розмістити пости, аби прогнозування даних було з найвищою точністю, що у свою чергу буде прияти прийняту рішень щодо покращення якості повітря у місті Вінниця [13].

Моніторинг та прогнозування якості повітря є перспективною сферою досліджень, оскільки все більше людей у всьому світі хвилюються

екологічними проблемами. Розробки в цій галузі можуть використовувати різні компанії та державні установи.

Наприклад, використовуючи системи моніторингу, організації можуть коригувати графік роботи працівників на відкритому повітрі, щоб уникнути негативного впливу на їхнє здоров'я. Державні органи можуть попереджати про небезпеку різних видів діяльності в дні з високим рівнем забруднення повітря, або планувати заходи в місті, виходячи з ситуації. Прогноз забруднення для забезпечення максимального комфорту населення. Крім того, на основі поточної інформації про якість повітря ви можете регулювати рівень шкідливих речовин, що викидаються в атмосферу, щоб уникнути перевищення норм забруднення [14].

На рисунку 1.5 зображена таблиця AQI – індексу якості повітря.

AQI	Категорія	Наслідки для здоров'я
0-50	Хороший	Мінімальний вплив
51-100	Задовільний	Може викликати незначний дискомфорт при диханні у чутливих людей. Не рекомендується займатись спортом на вулиці.
101-150	Помірно забруднений	Може спричинити дискомфорт при диханні у людей із захворюваннями дихальних шляхів, таких як астма, а також у людей з серцевими захворюваннями. Діти та літні люди зазнають найбільшого негативного впливу. Заняття спортом на вулиці заборонені.
151-200	Високий	Спричиняє дискомфорт при диханні у здорових людей. Людям із серцево-судинними і легневими захворюваннями категорично заборонено перебувати на вулиці. Заняття спортом на вулиці заборонені.
201-300	Дуже високий	Може викликати респіраторні захворювання у здорових людей при тривалому впливі. Людям із серцево-судинними і легневими захворюваннями категорично заборонено перебувати на вулиці. При такому забрудненні існує ризик інсульту та інфаркту у людей похилого віку. Заняття спортом на вулиці заборонені.
300+	Небезпечний	Впливає навіть на здорових людей, і може спричинити серйозні наслідки для їх здоров'я. Людям із серцево-судинними захворюваннями категорично заборонено перебувати на вулиці. Негативні наслідки можуть виникнути навіть під час легкої фізичної активності. Існує ризик інсульту та інфаркту.

Рисунок 1.5 – Таблиця AQI

Щоб громадськість могла зрозуміти поточну якість повітря, EPA (Агентство з охорони навколишнього середовища США) [10] встановило індекс якості повітря (AQI) [2], який вимірює п'ять найбільш поширених забруднювачів повітря. Індекс може надавати людям інформацію про

поточний стан повітря та інформувати про його можливий вплив на здоров'я людини.

Обчислюється AQI на основі концентрацій наступних забруднюючих речовин: приземного озону (O₃), оксиду вуглецю (CO), діоксиду сірки (SO₂), діоксиду азоту (NO₂) та завислих частинок (PM). Останній тип забруднювача можна розділити на два види: зважені частки розміром до 2,5 мкм та розміром від 2,5 до 10 мкм.

AQI може приймати значення від 0 до 500. Чим вище значення AQI, тим вище рівень забруднення повітря і тим більшу небезпеку для здоров'я повітря становить. Наприклад, значення AQI 50 говорить про те, що якість повітря є хорошим і ваш організм навряд чи зазнає негативного впливу, у той час як повітря зі значенням AQI більше 300 вже становить серйозну небезпеку для здоров'я людини [15].

Мета AQI – допомогти людям зрозуміти, якої якості повітря в даний момент і як перебування на відкритому повітрі може позначитися на здоров'ї людини. Для того, щоб це було легше зрозуміти, шкала AQI ділиться на шість категорій. Кожна категорія відповідає різному рівню якості повітря.

Для того, щоб люди могли швидше і простіше зрозуміти наскільки забруднене повітря, EPA надало кожній категорії AQI певний колір.

Наприклад, помаранчевий колір означає, що умови несприятливі для особливо чутливих, у той час як червоний колір означає, що умови можуть бути небезпечними для всіх, і так далі.

1.4 Особливості прогнозування рівня забрудненості

Першою складністю, з якою стикаються дослідники в цій галузі, є необхідність регулярного моніторингу якості повітря. Для якісного прогнозування необхідна мережа станцій моніторингу, щоб повністю охопити місто. Однак це не єдина проблема прогнозування якості повітря, цей процес складний з трьох причин.

1) Якість міського повітря залежить від багатьох складних факторів, таких як погодні умови, дорожній рух, промислові викиди, лісові пожежі тощо. Крім того, не всі ці дані оприлюднюються у відкритому доступі.

2) Якість міського повітря змінюється швидко і сильно. Це тому, що ступінь забруднення може змінюватися в залежності від різних факторів: від часу доби до годин пік, забруднення повітря є найбільш серйозним, погодні умови різко змінюються, бувають сильні вітри, зазвичай якість повітря набагато краще і багато інших факторів. Крім того, ступінь забруднення може відрізнятись від місця до місця. Наприклад, у промисловій зоні, коли компанія скидає забруднювачі, рівень забруднення може бути набагато вищим, ніж у житловому районі.

3) Якість повітря різко зміниться. Це може бути викликано екстремальними погодними умовами або надзвичайними ситуаціями. У деяких випадках, наприклад, сильний вітер або зливи, якість повітря може сильно змінитися за короткий проміжок часу [16].

Такі кардинальні відмінності важко передбачити, оскільки вони дуже рідкісні у всьому спостереженні, тому існуючі моделі не можуть добре передбачити раптові зміни.

1.5 Висновки

У даному розділі було здійснено аналіз предметної області, було визначено, що таке пости спостереження та які вони бувають, визначення моніторингу атмосферного повітря та його етапи. Також, розглянуто сучасний стан моніторингу атмосферного повітря у місті Вінниця, визначено основні проблеми та проекти, що займаються встановленням постів по місту.

2 ОСНОВНІ ЕТАПИ ВИКОНАННЯ РОБОТИ ТА ОГЛЯД НАБОРУ ВХІДНИХ ДАНИХ

2.1 Вибір оптимальних інформаційних технологій

З розвитком сфери інформаційних технологій, інформаційні системи стали обов'язковою частиною технічної бази більшості організацій. Використання передових інформаційних систем організації дозволяє поліпшити процеси її роботи. «Машинна праця» довела свою ефективність на практиці завдяки його можливості швидко і буквально обробляти дані.

На даний момент ринок запрошує велику кількість товарів та пропозицій. Для такого щоб організація благополучно вписувалася в рамки конкуренції, необхідно пристосувати всі її внутрішні процеси. Так як інформаційні системи є частиною організації і частиною цих процесів, то вони ще можуть зазнавати змін.

Головний елемент інформаційної системи – це інформація. З цієї підстави є потреба в розумінні того, як працювати з нею, для отримання на виході важливого результату. Оскільки інформаційна система – це програмний продукт, те й пристрій сприйняття інформації володітиме своєрідний образ [2]. Потрібно показати машині, як і яку інформацію обробляти. Для укладання даного завдання виділяють ряд критеріїв до наданої інформації:

Інформація має бути представлена в однорідному вигляді, тобто тримати спільну структуру. Насамперед варто відзначити, що інформація може обстановка надана в різних типах: наприклад, текстова, цифрова або змішана. Обумовлюючись певним завданням, необхідно для того, щоб система отримувала той чи інший певний образ даних.

Необхідно зрозуміти з яким обсягом даних доводиться працювати, наприклад як якщо інформації велика кількість, то і кількість інформаційного

шуму може бути досить величезним. Гігантські характеристики інформаційного шуму зводять нанівець необхідність роботи з даними.

Потрібно кваліфікувати значення даних, тобто зрозуміти як дані взаємопов'язані між собою. У випадку, коли інформація не має взаємозв'язків, знижується ефективність роботи з такою інформацією і ця робота може не призвести до необхідних результатів.

Інформація повинна мати користь. Це означає, що, проаналізувавши інформацію, висновки призведуть до певного значення і використовуватимуться практично.

Є належне сучасні та актуальні способи аналізу даних в інформаційних системах:

- штучного походження нейронні мережі (ІНС);
- факторний аналіз;
- регресійний аналіз;
- тест панельних даних;
- кореляційний аналіз;
- семантичний аналіз;
- Data Mining [4].

Перш ніж говорити про вибір методу даних, слід відзначити фундаментальну основу будь-якого програмного продукту - основне завдання або завдання, яку буде вирішувати даний товар.

Від спрямованості залежить, яку інформацію необхідно давати на вхід інформаційної системи, а від того який вона стане на вході вже залежить сам вибір методу. Таким чином, якщо вектор містить, наприклад, фінансове спрямування, то на вхід буде сервіруватися фінансова інформація (частіше втілена в числах), а як метод буде застосовуватися метод, що працює з такою інформацією.

Найбільш часто в організаціях використовуються фінансові інформаційні системи, тобто системи, що працюють з фінансовою інформацією і на виході дають відповідного сімейства результати. Під

фінансовою інформацією розуміються відомості про процеси виготовлення, речових ресурсів, трудових ресурсів, грошові процеси та про стан об'єктів управління на конкретний час. Вона у процесі виробничої домашньої роботи суб'єктів господарювання, та її призначення вироблено у використанні для управлінських справ. Особливості фінансової інформації призначають потребу та фінансову необхідність використання спеціальних технічних засобів, у яких кількості засобів обчислювальної техніки, при її збиранні, накопиченні, передачі та обробці. Тому фінансова інформація вважається як предметом, наприклад і продуктом автоматичної обробки. Цей малюнок інформації найчастіше виявляється у числовому форматі, тому методам його обробки необхідно шукати причинно-наслідкові зв'язку в числових структурах для цілей можуть використовуватися методи факторного аналізу, регресійного аналізу та аналізу панельних даних [3].

Інформація, подана на вхід інформаційної системи може бути представлена в змішаному вигляді або мати відфільтрований вигляд, тобто однорідну числову або текстову структуру. У випадку, коли інформація виявлена в змішаному вигляді, в залежності від цілей її приводять до однорідного вигляду за допомогою методу Data mining або вручну, або користуються методами, здатними без участі людини обробляти цю інформацію – штучного походження нейронні мережі або недійсний розум [12].

Завдання з обробки змішаної інформації можуть поєднуватися до «змішування» або комбінування методів, в залежності від встановлених завдань. Таким чином, відфільтрувавши інформацію за допомогою методу Data Mining і отримавши однорідну числову структуру, можна застосувати метод факторного аналізу.

Повернувшись до типізації інформації варто позначити таку картину як текстова інформація. Обробляючи такого сімейства інформацію, застосовуються методи семантичного аналізу. Ці методи є важким математичним завданням для «машини» і зазвичай застосовуються у процесі

створення штучного походження розуму, хоча вважаються ефективним помічником для інформаційних систем сфер торгівлі, маркетингу, журналістики тощо.

Таким чином, перед тим як обирати метод для аналізу даних в інформаційних системах, необхідно визначитися з метою або завданням, яке буде виконуватися за допомогою програмного продукту. Після того як виявлено завдання, необхідно відзначити ряд критеріїв до вхідної інформації, а потім привести її до важливого вигляду. І тільки після того як інформація містить важливий вигляд, розробник програмного забезпечення переходить до кроку використання методу аналізу даних [17].

2.1.1 ArcGIS Online

Для вирішення поставленого завдання було обрано інтернет середовище ArcGIS Online (рис. 2.1), дане середовище використовується для реалізації функцій картографування та аналізу. Дане середовище дозволяє створювати власні карти, виконувати аналіз даних.

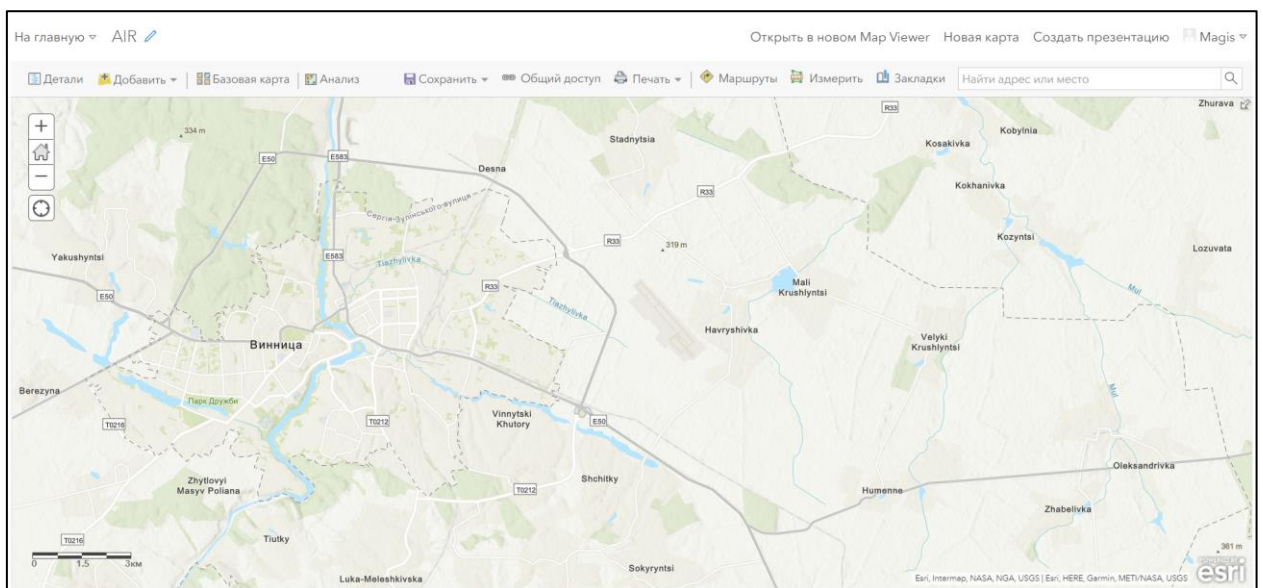


Рисунок 2.1 – Середовище ArcGIS Online

ArcGIS є повною системою, яка дозволяє збирати, організовувати, управляти, аналізувати, обмінюватися та розподіляти географічну інформацію. Як світовий лідер серед платформ для побудови та використання геоінформаційних систем (ГІС), ArcGIS використовується людьми по всьому світу для застосування географічних знань у практичній сфері державного управління, бізнесу, науки, освіти та ЗМІ. Платформа ArcGIS дозволяє публікувати географічну інформацію для доступу та використання будь-якими користувачами. Система доступна в будь-якій точці, де можливе використання веб-браузерів, мобільних пристроїв у вигляді смартфонів та настільних комп'ютерів.

ArcGIS Online містить інтерактивні карти та 3D сцени, що дозволяє переглядати, вивчати та аналізувати географічні дані. Використовувати ArcGIS Living Atlas of the World, динамічну колекцію карт, сцен, шарів даних, зображень, аналітику та програми із спільноти ArcGIS, інструменти аналізу, які містить Map Viewer Classic, для пошуку нових закономірностей, відповідних місць, збагачення даних, визначення найближчих об'єктів та підсумовування даних.

Модуль Geostatistical Analyst надає найефективніші інструменти як інтерполяції даних, так статистичного аналізу модельних поверхонь [18].

Коли дані представлені не в повному обсязі або містять помилки, ArcGIS Geostatistical Analyst надає ймовірну основу для кількісної оцінки невизначеностей та виключення помилок із прогнозу. Створення поверхонь на основі вимірювань з використанням таких методів інтерполяції:

- метод зворотних завислих відстаней;
- радіальні базисні функції;
- глобальні та локальні поліноми;
- кригінг для точних даних та даних з помилками;
- кокригінг (багатомірна версія вищезгаданих моделей кригінгу);
- ізотропні чи анізотропні моделі.

Створено веб-сервіси (<https://arccg.is/1XTiDS>) для автоматизації процесу оцінювання параметрів джерел понаднормативних викидів на основі нечіткої бази знань з візуалізацією результатів у Google Maps з використанням запропонованого у роботі методу та інформаційної технології.

Створені веб-сервіси вільно доступні в Інтернеті, що дозволяє науковцям, викладачам, фахівцям та студентам використовувати ці засоби для наукових досліджень та у навчальних цілях.

2.1.2 Kriging інтерполяція

Крігінг – це метод інтерполяції, який робить прогнози в місцях без вибірки, використовуючи лінійну комбінацію спостережень у найближчих вибіркових місцях. Вплив кожного спостереження на прогноз крігінгу заснований на кількох факторах:

- 1) його географічна близькість до місця без вибірки,
- 2) просторове розташування всіх спостережень (тобто конфігурація даних, наприклад групування спостережень у районах із надмірною вибіркою),
- 3) закономірність просторової кореляції даних [19].

Розробка моделей крігінга має сенс лише тоді, коли дані просторово корелюють. Крігінг має кілька переваг перед традиційними методами інтерполяції, такими як зважування зворотної відстані або найближчий сусід:

- 1) він забезпечує міру невизначеності, пов'язану з результатами (тобто дисперсію крігінгу);
- 2) він враховує залежні від напрямку відносини (тобто просторова анізотропія);
- 3) ваги призначаються спостереженням на основі просторової кореляції даних замість припущень, зроблених аналітиком для IDW;
- 4) передбачення крігінга не обмежені діапазоном спостережень, що використовуються для інтерполяції,

5) дані, виміряні на різних просторових опорах, можна об'єднати і змінити підтримку, наприклад, зменшення чи збільшення.

Метод крігінга походить із галузі геостатистики. Етимологічно термін «геостатистика» позначає статистичне вивчення природних явищ. Перші розробки геостатистики в 1950-х і 1960-х роках мали на меті покращити оцінку видобуваних запасів у гірничих родовищах (Kriging, 1951; Journel and Huibregts, 1978). Сфера його застосування значно розширилася, щоб охопити сьогодні більшість галузей геонауки (наприклад, геологію, геохімію, геогідрологію, ґрунтознавство) і широкий спектр дисциплін, які займаються аналізом даних простору-часу, таких як океанографія, гідрологія, дистанційне зондування, сільського господарства, географії, ГІС та природокористування [18].

Більшість застосувань геостатистики пов'язані з прогнозуванням вимірних атрибутів у місцях без вибірки. Розглянемо, наприклад, передбачення значення атрибута z в місці u_0 за допомогою даних у п'яти місцях, u_1 до u_5 (рис. 2.2).

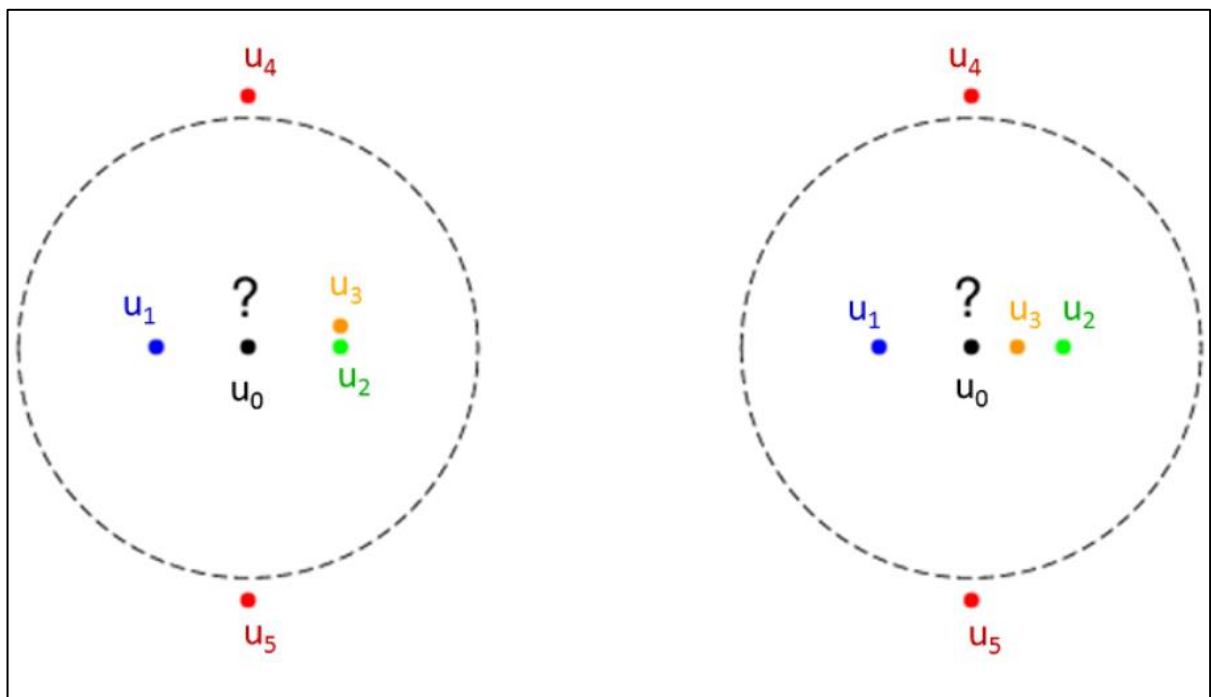


Рисунок 2.2 – Простий приклад крігінга

Дві двовимірні конфігурації даних з різними положеннями для розташування u_3 . В обох випадках передбачення в місці u_0 виконується з використанням звичайного крігінга та п'яти даних у розташуваннях $u_1 - u_5$. Радіус пунктирного кола з центром u_0 відповідає діапазону кореляції.

2.1.3 Kaggle

Для вирішення поставленого завдання було запропоновано, скористатись інтернет середовищем Kaggle.

Kaggle – це платформа для змагань, де вирішуються реальні задачі системного аналізу, передбачуваного моделювання, в рамках якого статистики та добувачі даних конкурують у створенні найкращих моделей. Чому для вирішення поставленої задачі було обрано саме Kaggle? По-перше, Kaggle пропонує цікаві задачі та складні проекти, які допомагають розвивати свої здібності. По-друге, можливість працювати в команді, або ж завести нових друзів, спілкування як з досвідченими користувачами, так і з новачками. І останнє, це можливість потрапити до найбільшої спільноти з даними у світі [2].

На рисунку 2.3 зображено датасет, на основі якого було виконано дану роботу.

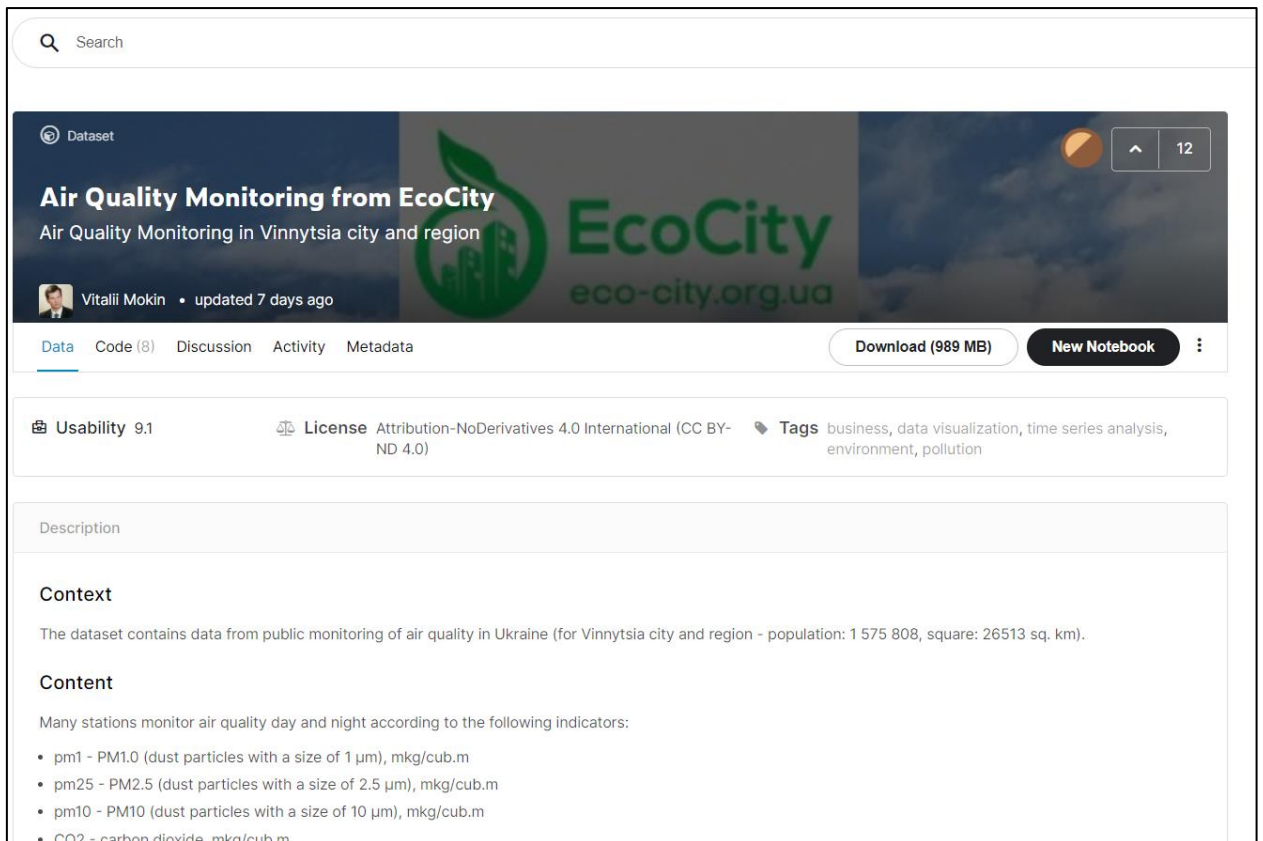


Рисунок 2.3 – Датасет у середовищі Kaggle

Для розв’язання головного завдання було виконано геостатистичний аналіз за допомогою SciKit-GStat.

SciKit-Gstat – це модуль аналізу в стилі scipy для аналізу варіограм. Базовий клас називається Variogram, що, ймовірно, є єдиним необхідним імпортом. Однак існує кілька інших класів:

- DirectionalVariogram для спрямованих варіограм;
- SpaceTimeVariogram для просторово-часових варіограм;
- ЗвичайнийКрігінг для інтерполяції;
- MetricSpace для попередньо обчислених просторових вибірок [5].

Класи варіограм мають подібний інтерфейс і можуть обчислювати експериментальні варіограми та відповідати теоретичним функціям моделі варіограм. Модуль використовує багатий вибір оцінок напівдисперсій, функцій моделі варіограми та функцій spatial binning, водночас будучи розширюваним.

У геостатистиці процедура просторової інтерполяції відома як Крігінг. Це сходить до винахідника Крігінга, південноафриканського гірничого інженера на ім'я Дейв Крідж. Він опублікував метод у 1951 році. У багатьох підручниках є термін передбачення, але Крігінг все ще базується на припущенні, що змінна є випадковим полем. Тому я віддаю перевагу терміну оцінка і позначу метод Крігінга BLUE, Best Linear Unbiased Estimator. Загалом, мета полягає в тому, щоб оцінити змінну в місці, яке не спостерігалось, використовуючи спостереження з близьких місць. Крігінг вважається найкращою оцінкою, оскільки ми використовуємо просторову структуру, описану варіограмою, щоб знайти відповідні вагові показники для усереднення спостережень у близьких місцях [7].

Враховуючи набір точок спостереження s_i і значень спостереження в цих місцях $Z(s_i)$, вже можна стверджувати, що оцінка в неспостережуваному місці $Z^*(s_0)$ є середнім зваженим, формула (2.1):

$$\sum_{i=0}^N \lambda_i * Z(s_i) = Z * (s_0), \quad (2.1)$$

де N – розмір s , а λ – масив ваг. Це те, що ми хочемо розрахувати з підібраної моделі варіограми.

Замість того, щоб просто складати ваги, ми тепер дізнаємося, як ми можемо використовувати модель варіограми для обчислення ваг. По суті варіограма описує, як точкові спостереження стають більш несхожими з відстанню. Відстань точок можна легко розрахувати не тільки для спостережуваних, а й неспостережуваних місць. Оскільки варіограма є лише функцією відстані, ми можемо легко обчислити значення напівдисперсії для будь-якої можливої комбінації пар точок.

Ми наблизилися до алгоритму Крігінга. Єдине, чого не вистачає, це гарантувати неупередженість. Вага сягає майже одиниці, але вони не єдині. Ми хочемо переконатися, що вони завжди єдині. Це робиться шляхом

додавання ще одного рівняння до системи лінійних рівнянь. Крім того, ми перейменуємо масив a на λ , який частіше використовується для ваг Кригінга.

Рівняння Кригінга для звичайного Кригінга, яке можна знайти в підручниках. Ми додали одиниці до масиву результатів і до матриці напівваріацій. μ – це множник Лагранжа, який буде використовуватися для оцінки дисперсії Кригінга, про яку ми розглянемо пізніше. Звичайний Кригінг все ще передбачає, що спостереження та їх залишки є нормально розподіленими та стаціонарністю другого порядку.

Просторова інтерполяція – це ефективний спосіб дезагрегувати дані або зменшити їх до нижчого рівня шляхом їхнього виведення [17]. У цьому посібнику ми розглянули лише один варіант використання, в якому інтерполювали дані про населення, які не були доступні в адміністративних районах нижнього рівня.

2.1.4 Приклади застосування інтерполяції

Інтерполяція передбачає значення для комірок у растрі з обмеженої кількості вибірових точок даних. Його можна використовувати для прогнозування невідомих значень для будь-яких даних географічної точки, таких як висота над рівнем моря, кількість опадів, концентрація хімічних речовин і рівень шуму.

Вхідними даними тут є точковий набір відомих значень рівня опадів, показаний на ілюстрації зліва. На ілюстрації праворуч показано растр, інтерпольований з цих точок [2]. Невідомі значення прогноуються за допомогою математичної формули, яка використовує значення найближчих відомих точок (рис. 2.4).

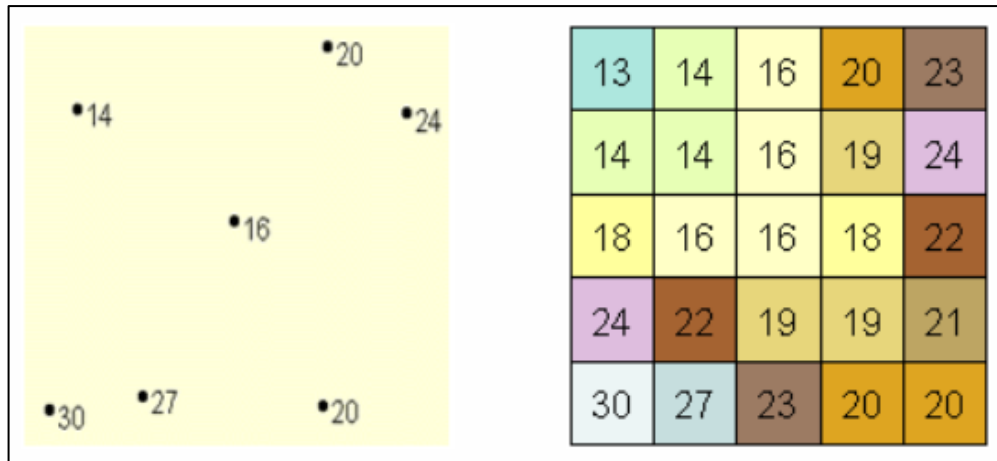


Рисунок 2.4 – Інтерполяція дощової поверхні

Типове використання точкової інтерполяції – створення поверхні висот із набору вимірів зразка.

На наступному графіку кожен символ у точковому шарі представляє місце, де було виміряно висоту (рис. 2.5). За допомогою інтерполяції будуть передбачені значення для кожної клітинки між цими вхідними точками.

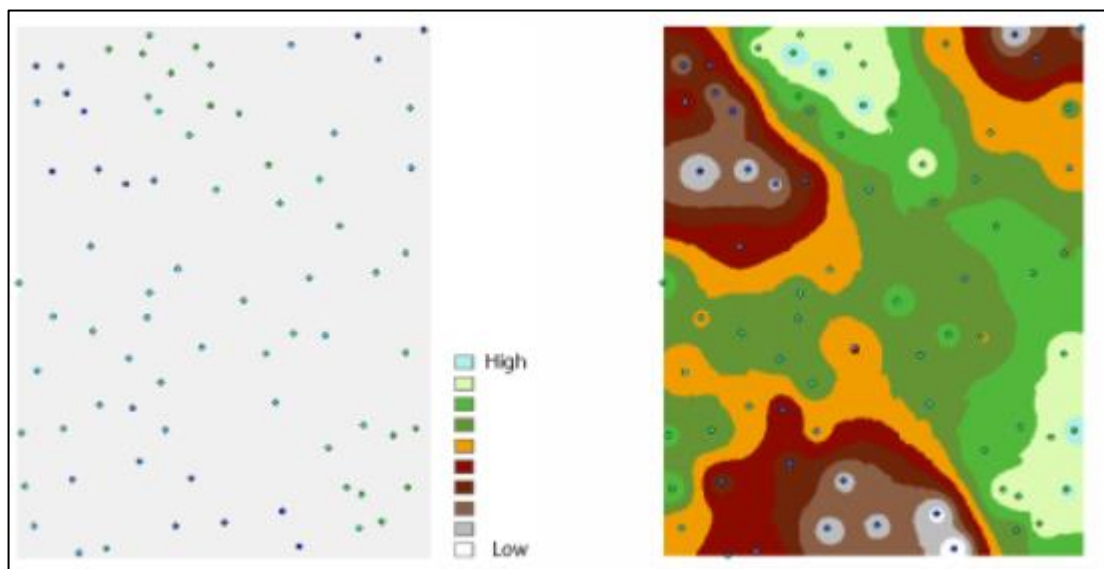


Рисунок 2.5 – Інтерполяція поверхні висот

У наведеному нижче прикладі інструменти інтерполяції були використані для вивчення кореляції концентрації озону з захворюваннями легенів у Каліфорнії. На зображенні зліва показано розташування станцій

моніторингу озону [3]. Зображення праворуч відображає інтерпольовану поверхню, забезпечуючи прогнози для кожного місця в Каліфорнії. Поверхня була отримана за допомогою крігінга (рис. 2.6).

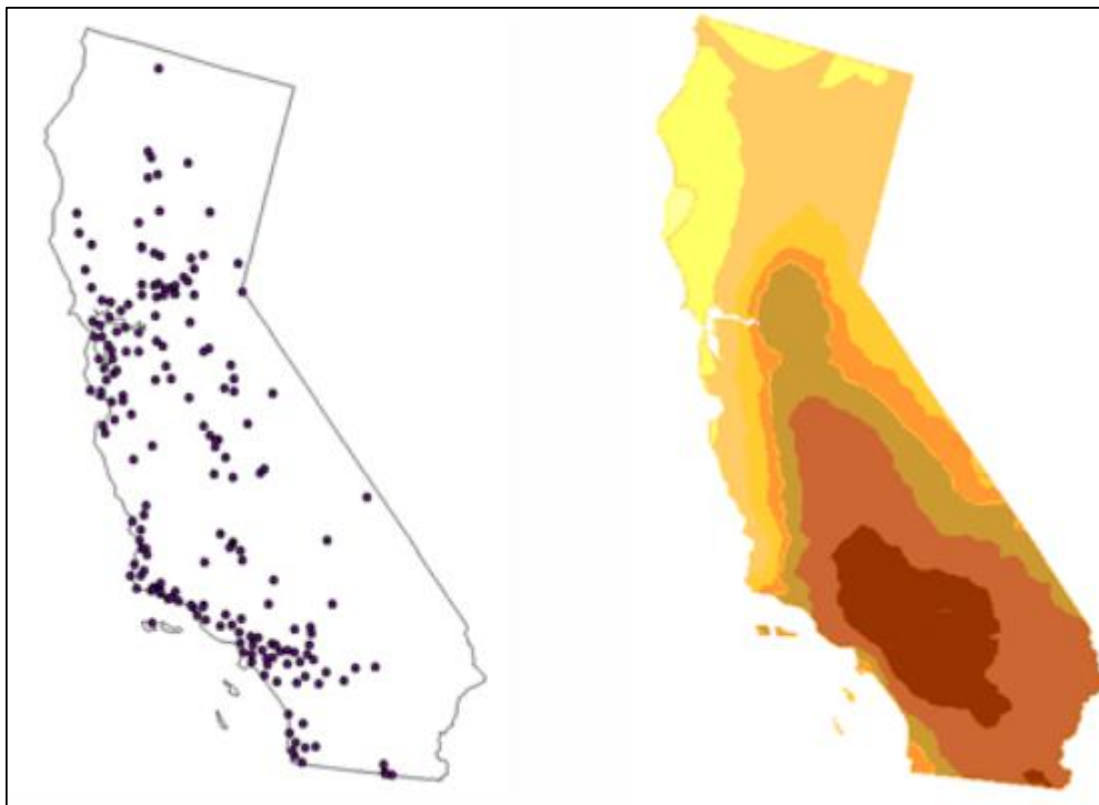


Рисунок 2.6 – Інтерполяція поверхні концентрації

2.1.5 Інструменти інтерполяції

Інструменти інтерполяції поверхні створюють безперервну (або прогнозовану) поверхню з вибірових значень точки.

Відвідування кожного місця в зоні дослідження для вимірювання висоти, концентрації або величини явища, як правило, є складним або дорогим. Замість цього ви можете виміряти явище в стратегічно розосереджених місцях вибірки, а прогнозовані значення можна призначити всім іншим місцям. Вхідні точки можуть бути як випадковим, так і регулярним, або заснованими на схемі вибірки [2].

Неперервне представлення поверхні набору растрових даних являє собою певну міру, таку як висота, концентрація або величина (наприклад, висота, кислотність або рівень шуму). Інструменти поверхневої інтерполяції роблять прогнози на основі вибіркового вимірювання для всіх місць у вихідному наборі растрових даних, незалежно від того, було проведено вимірювання в місці розташування чи ні.

Існує безліч способів отримати прогноз для кожного місця; кожен метод називається моделлю. З кожною моделлю існують різні припущення щодо даних, і певні моделі більш застосовні для конкретних даних — наприклад, одна модель може врахувати місцеві варіації краще, ніж інша. Кожна модель створює прогнози за допомогою різних розрахунків.

Інструменти інтерполяції загалом поділяються на детерміновані та геостатистичні методи [18].

Методи детермінованої інтерполяції призначають значення місцям на основі навколишніх вимірюваних значень і на основі заданих математичних формул, які визначають гладкість результуючої поверхні.

До детермінованих методів належать IDW (зважування зворотної відстані), Natural Neighbor, Trend і Spline.

Геостатистичні методи базуються на статистичних моделях, які включають автокореляцію (статистичний зв'язок між виміряними точками). Через це геостатистичні методи не лише здатні створити поверхню прогнозу, але й забезпечують певну міру впевненості чи точності передбачень.

Крігінг є геостатистичним методом інтерполяції.

Інші інструменти інтерполяції, Tопо to Raster і Tопо to Raster by File, використовують метод інтерполяції, спеціально розроблений для створення безперервних поверхонь із контурних ліній, а також ці методи містять властивості, сприятливі для створення поверхонь для гідрологічного аналізу.

2.1.6 Порівняння методів інтерполяції

Інтерполяція передбачає значення для комірок у растрі з обмеженої кількості вибірових точок даних. Його можна використовувати для передбачення невідомих значень для будь-яких даних географічної точки, таких як висота над рівнем моря, кількість опадів, концентрація хімічних речовин, рівень шуму тощо.

Нижче наведено доступні методи інтерполяції.

Інструмент IDW (Inverse Distance Weighted) використовує метод інтерполяції, який оцінює значення комірки шляхом усереднення значень вибірових точок даних в околиці кожної комірки обробки. Чим ближче точка до центру оцінюваної комірки, тим більший вплив або вагу вона має на процес усереднення (рис. 2.7).

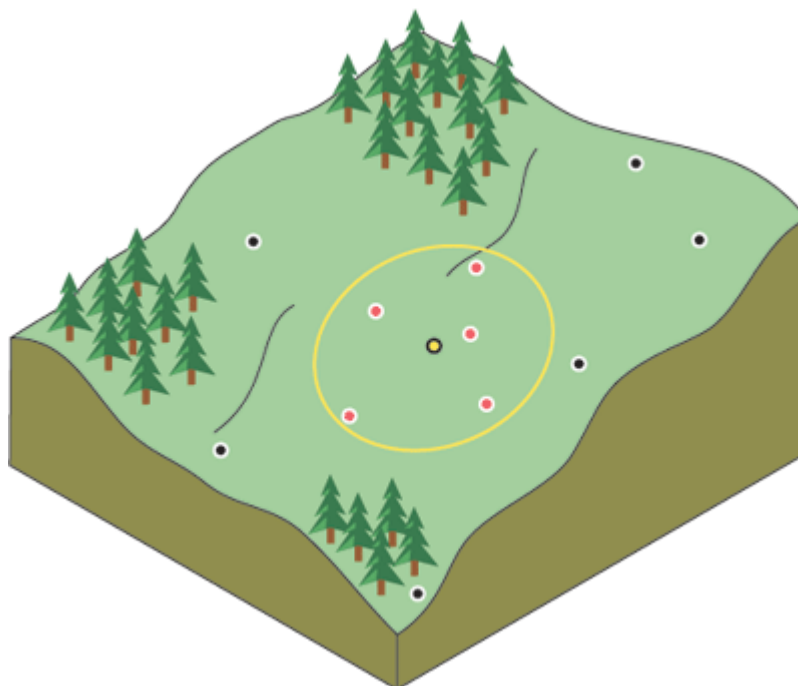


Рисунок 2.7 – Метод інтерполяції IDW

Крігінг – це вдосконалена геостатистична процедура, яка генерує оцінку поверхні з розсіяного набору точок із значеннями z (рис. 2.8). Більше, ніж інші методи інтерполяції, перед вибором найкращого методу оцінки для

генерування вихідної поверхні слід провести ретельне дослідження просторової поведінки явища, представленого z-значеннями [14].

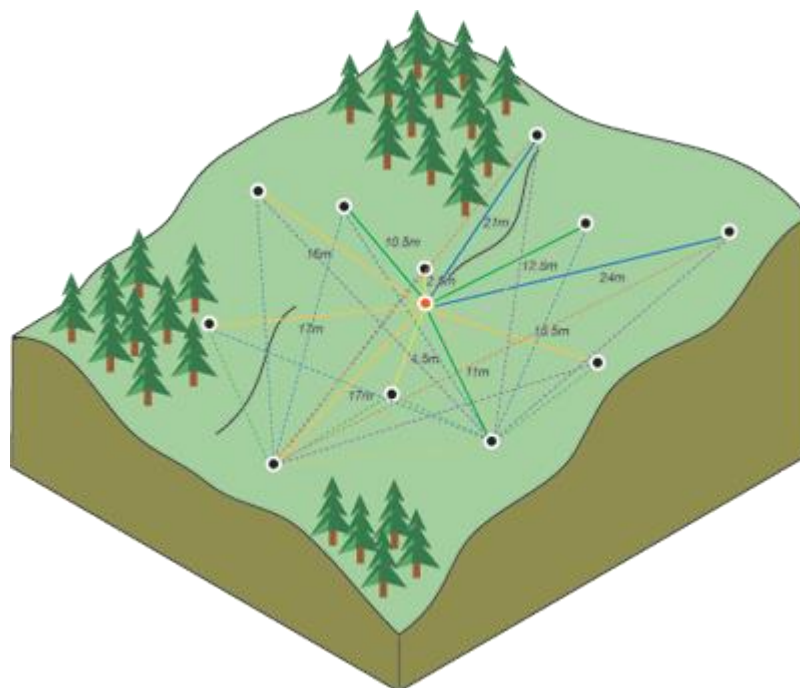


Рисунок 2.8 – Метод інтерполяції Крігінг

Інтерполяція Natural Neighbor знаходить найближчу підмножину вхідних вибірок до точки запиту та застосовує до них ваги на основі пропорційних площ для інтерполяції значення (Сібсон, 1981) [2]. Вона також відома як інтерполяція Сібсона або інтерполяція, що викрадає область (рис 2.9).

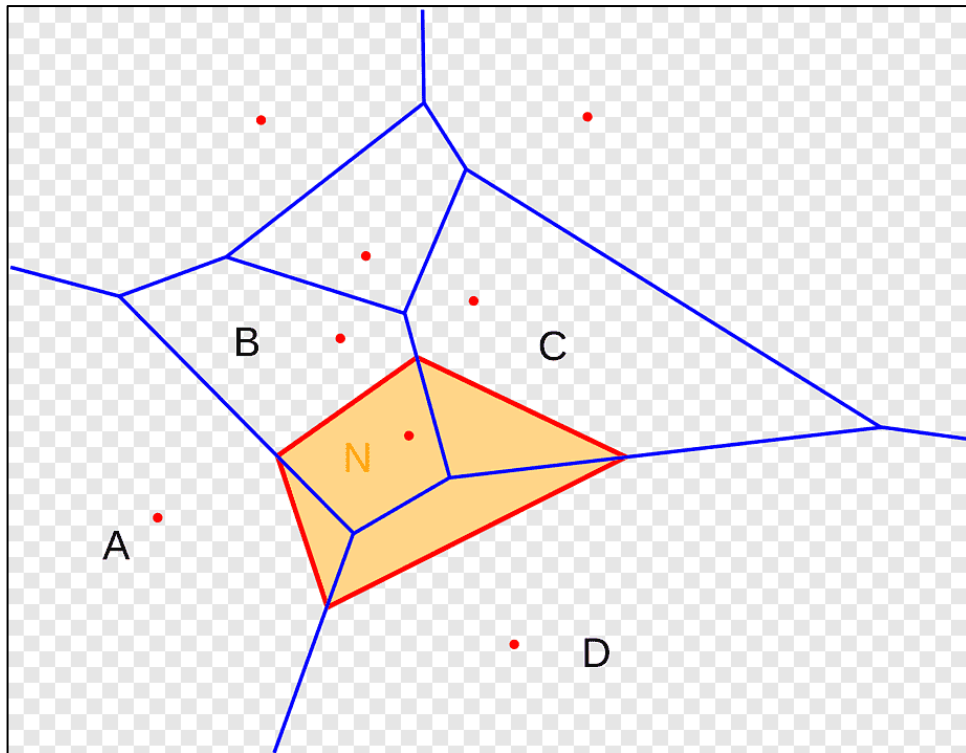


Рисунок 2.9 – Інтерполяція Natural Neighbor

Інструмент Сплайн використовує метод інтерполяції, який оцінює значення за допомогою математичної функції, яка мінімізує загальну кривизну поверхні, що призводить до гладкої поверхні, яка проходить точно через вхідні точки (рис. 2.10) [11].

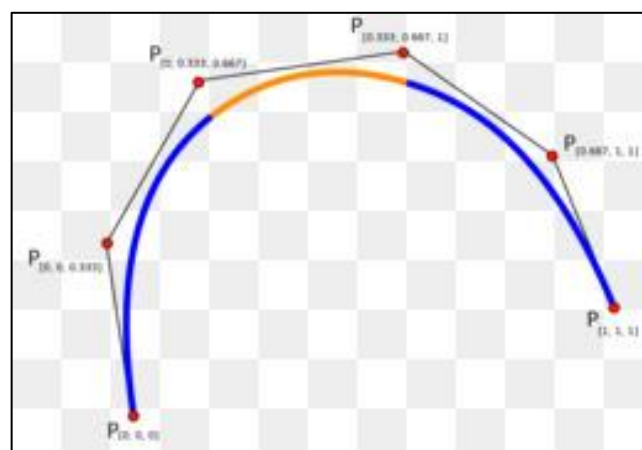


Рисунок 2.10 – Сплайн інтерполяція

Інструмент «Сплайн із бар'єрами» використовує метод, подібний до техніки, що використовується в інструменті «Сплайн», з основною відмінністю в тому, що цей інструмент враховує розриви, закодовані як у вхідних бар'єрах, так і в даних точки введення (рис. 2.11).

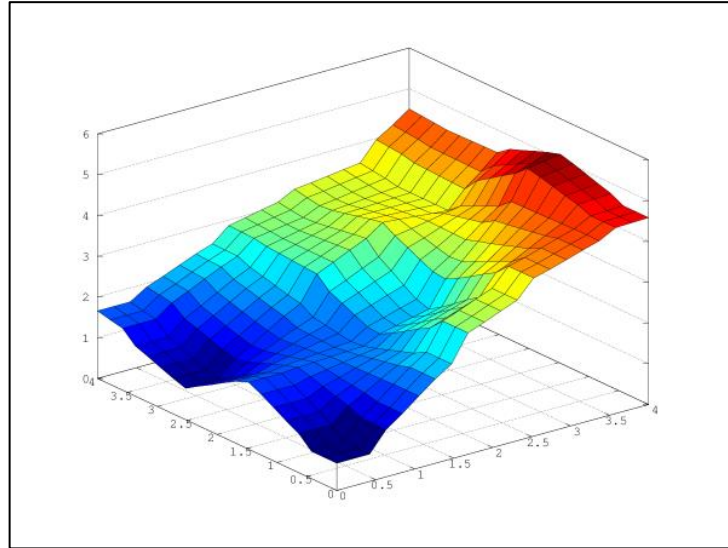


Рисунок 2.11 – Інструмент «Сплайн із бар'єрами»

Інструменти Topo to Raster і Topo to Raster by File використовують техніку інтерполяції, спеціально розроблену для створення поверхні, яка більш точно представляє природну дренажну поверхню і краще зберігає лінії хребтів і потокові мережі від вхідних даних контуру (рис. 2.12) [6].

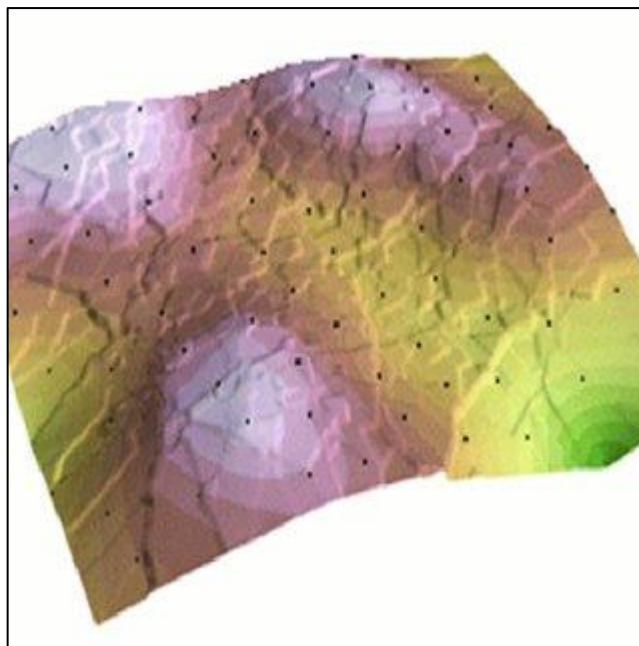


Рисунок 2.12 – Інструменти Topo to Raster

Використовуваний алгоритм базується на алгоритмі ANUDEM, розробленому Хатчинсоном та іншими в Австралійському національному університеті [2].

Тренд – це глобальна поліноміальна інтерполяція, яка відповідає гладкій поверхні, визначеній математичною функцією (поліномом), до точок вхідної вибірки. Поверхня тенденції поступово змінюється і фіксує грубомасштабні моделі даних (рис. 2.13).

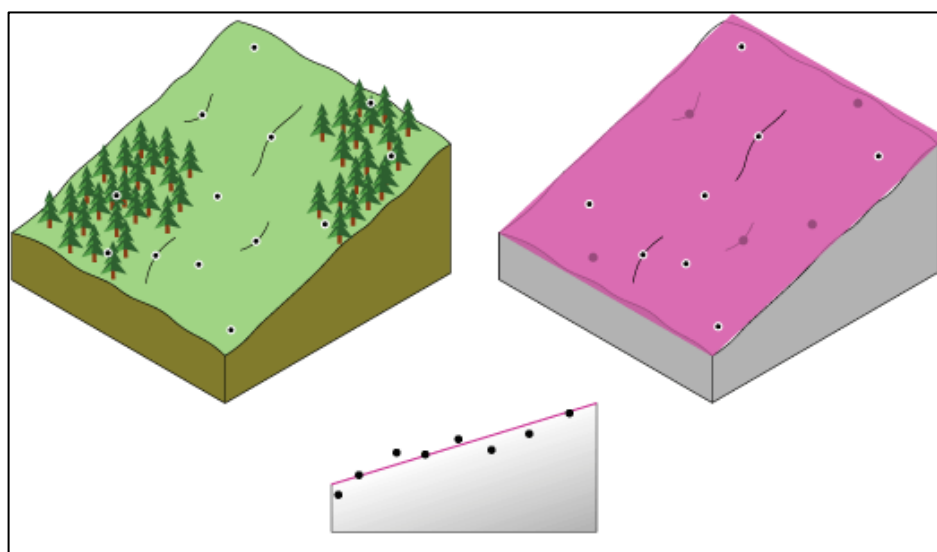


Рисунок 2.13 – Тренд інтерполяція

2.2 Огляд вхідного набору даних

Однією з найважливіших навичок, для розв'язання задач, є вміння правильно вивчати дані. Ретельний розвідувальний аналіз даних має важливе значення для забезпечення цілісності зібраних даних та проведеного аналізу. Дослідження даних за допомогою добре побудованих візуалізацій та описової статистики – це чудовий спосіб ознайомитися з даними, та сформулювати гіпотези щодо своїх спостережень. Завдяки цьому процесу чи можемо визначити або підтвердити постановку проблеми, що дуже важливо [4].

Цілі розвідувального аналізу даних:

- 1) Опис даних, перевірка на відсутність значень, попередній перегляд, виконання стандартних функцій на знаходження count, mean, std, квантилії та інші;
- 2) Очищення пошкоджених даних, обробка відсутніх даних, виявлення аномальних даних та об'єднання їх під одну категорію;
- 3) Візуалізація розподілу даних, побудова діаграм, гістограм, графіків, боксплотів;
- 4) Побудова матриці кореляції, теплової карти, візуалізація простору ознак за допомогою моделі зменшення розмірності PCA, за необхідності [7].

Як було вказано у підрозділі 1.2, у Вінниці наразі встановлено 9 постів моніторингу якості атмосфери повітря. Тож для розв'язання даної задачі ресурсами Eco City та SaveEcoBot було надано 9 наборів даних по цих постах. У таблиці 2.1 наведені стовпці з id_station – унікальний номер станції, locality – назвою тої чи іншої станції, address – адресою станції та source – сервіс який надав дані по тій чи іншій станції.

Таблиця 2.1 – Вхідний набір даних

id_station	locality	address	source
337	Олієжир	Чехова, 21	Eco City
767	Соборна 36	Соборна, 36	Eco City
774	Станція Замостя	Замостянська, 7/35	Eco City
790	Станція Вишенька	Космонавтів, 30	Eco City
848	Верхня Слов'янка	Константиновича, 4	Eco City
1315	ВНТУ	Хмельницьке шосе, 95 (буд. 7)	Eco City
1183	Авалон	Трамвайна, 3	Eco City
761	Тріумф	600-річчя, 66в	Eco City
1008	Сеценка 75	Сеценка, 75	SaveEcoBot

Як бачимо з таблиці (2.1) було взято дані по 8 станціях з Eco City, та 1 станція з SaveEcoBot. За адресами вказаними у таблиці (2.1) було взято координати точок та нанесено на карту в ArcGIS Online (рис. 2.14).

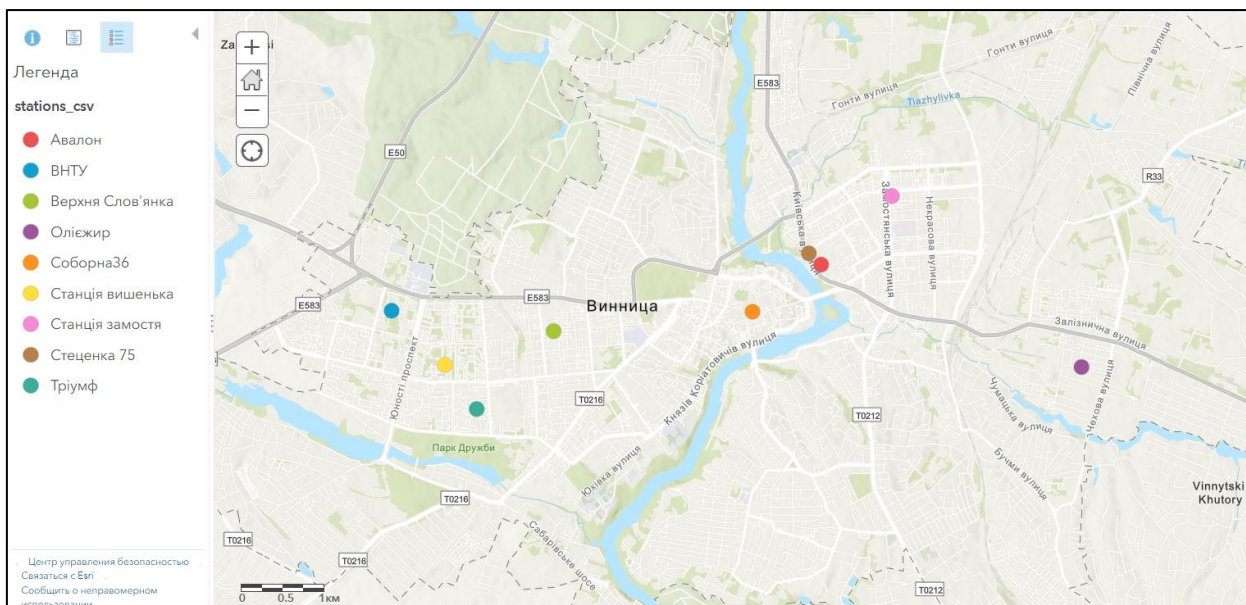


Рисунок 2.14 – Карта з координатами точок

Можна робити висновок, що станції розміщені майже в лінію, а не розкидані по усьому місту.

Після того, як дані були оброблені в MS Access (рис 2.15) їх було прикріплено по полю `id_station` до постів на карту ArcGIS Online.

id_station	datetime	CO2	Humidity	PM1_0	PM10	PM2_5	Pressure	Temperature
1183	11.01.2021 09:29	550.0	32.48	69.11	128.11	113.0	99265.84	24.11
1183	11.01.2021 09:36	550.0	32.75	70.0	125.9	112.8	99273.22	25.96
1183	11.01.2021 09:41	550.0	32.75	67.9	124.9	111.2	99263.88	26.76
1183	11.01.2021 09:42	550.0	44.64	63.4	116.8	106.7	99281.64	25.14
1183	11.01.2021 10:22	550.0	46.47	62.5	109.8	99.3	99294.29	21.09
1183	11.01.2021 10:23	550.0	43.37	59.33	109.78	100.89	99290.63	21.1
1183	11.01.2021 10:24	1451.0	43.03	61.5	112.5	100.7	99296.45	21.29
1183	11.01.2021 10:26	1428.0	43.1	60.4	106.6	96.0	99292.97	21.43
1183	11.01.2021 10:27	1407.0	42.81	60.73	108.0	99.36	99295.47	21.59
1183	11.01.2021 10:28	1396.0	42.59	60.2	109.4	96.5	99293.63	21.75
1183	11.01.2021 10:29	1375.0	42.24	63.27	109.27	98.18	99293.27	21.93
1183	11.01.2021 10:30	1447.0	42.67	58.56	108.89	95.78	99287.98	22.14
1183	11.01.2021 10:42	550.0	44.37	57.8	100.8	89.5	99298.53	22.59
1183	11.01.2021 10:43	550.0	44.25	57.7	102.1	89.8	99300.92	22.32
1183	11.01.2021 10:44	2456.0	44.82	54.33	105.11	93.44	99296.34	22.26
1183	11.01.2021 10:45	2286.0	46.73	56.11	99.56	89.11	99299.64	22.28
1183	11.01.2021 10:47	1950.0	43.19	55.7	93.0	86.4	99298.34	22.24
1183	11.01.2021 10:48	1819.0	43.6	56.2	99.8	88.8	99295.59	22.17

Рисунок 2.15 – Обробка даних в MS Access

На рисунку 2.15, бачимо завершений варіант таблиці, де дані було транспоновано та відсортовано по ознаках. Загалом усі станції фіксували значення за такими показниками:

- pm1 – PM1.0 (частинки пилу розміром 1 мкм), мкг/куб.м;
- pm25 – PM2,5 (частинки пилу розміром 2,5 мкм), мкг/куб.м;
- pm10 – PM10 (частинки пилу розміром 10 мкм), мкг/куб.м;
- CO2 – вуглекислий газ, мкг/куб.м;
- temperature – температура, в градусах Цельсія;
- вологість – Вологість, %;
- тиск – Тиск, міліметри ртутного стовпа.

Однак, не усі станції вимірювали за усіма, перерахованими вище, показниками. У середовище ArcGIS Online було додано усі таблиці з бази даних та прикріплено до точок по полю id_station (рис. 2.16).

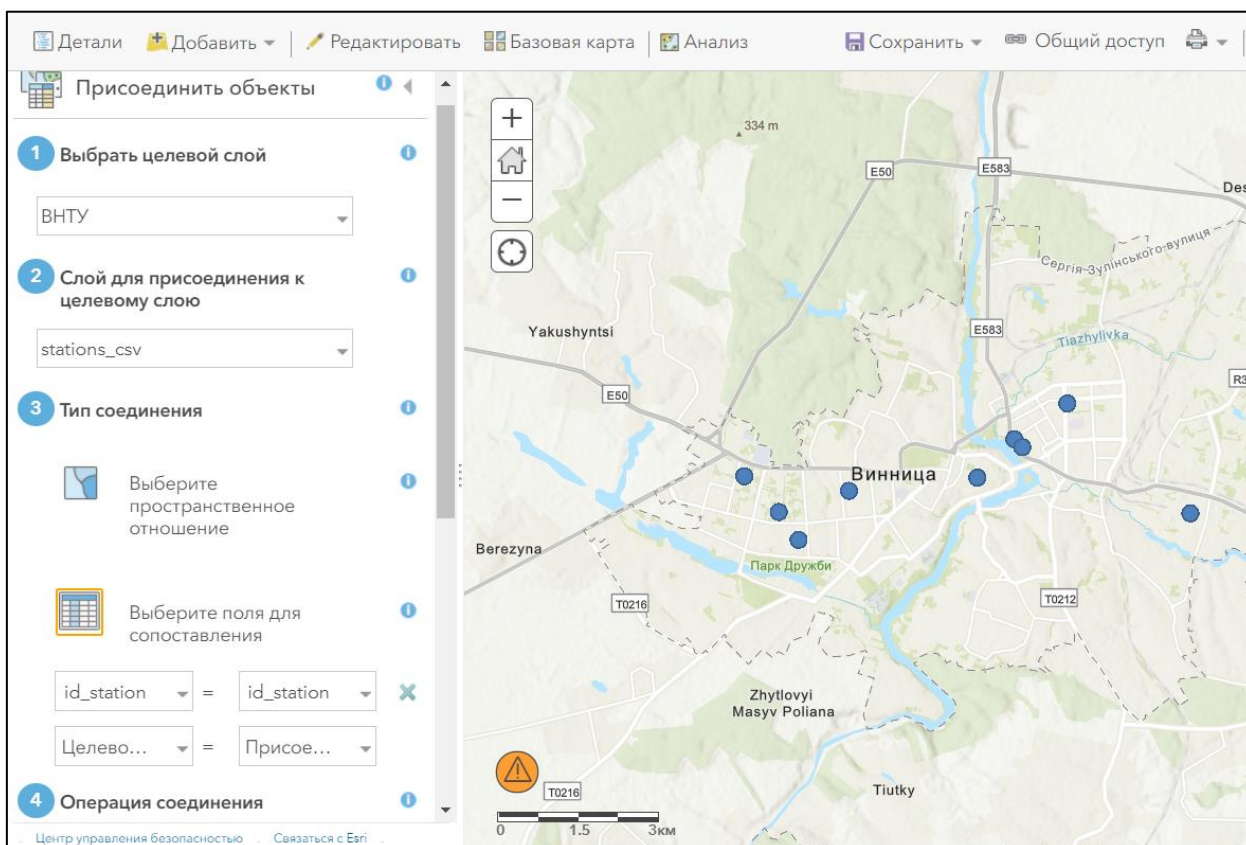


Рисунок 2.16 – Прив'язування набору даних у таблиці ВНТУ до карти з розміченими постами

Розглянемо кожну станцію окремо, після накладання шарів усі точки відображаються однаковим кольором. На рисунку 2.17 було зображено карту станції Олієжир.

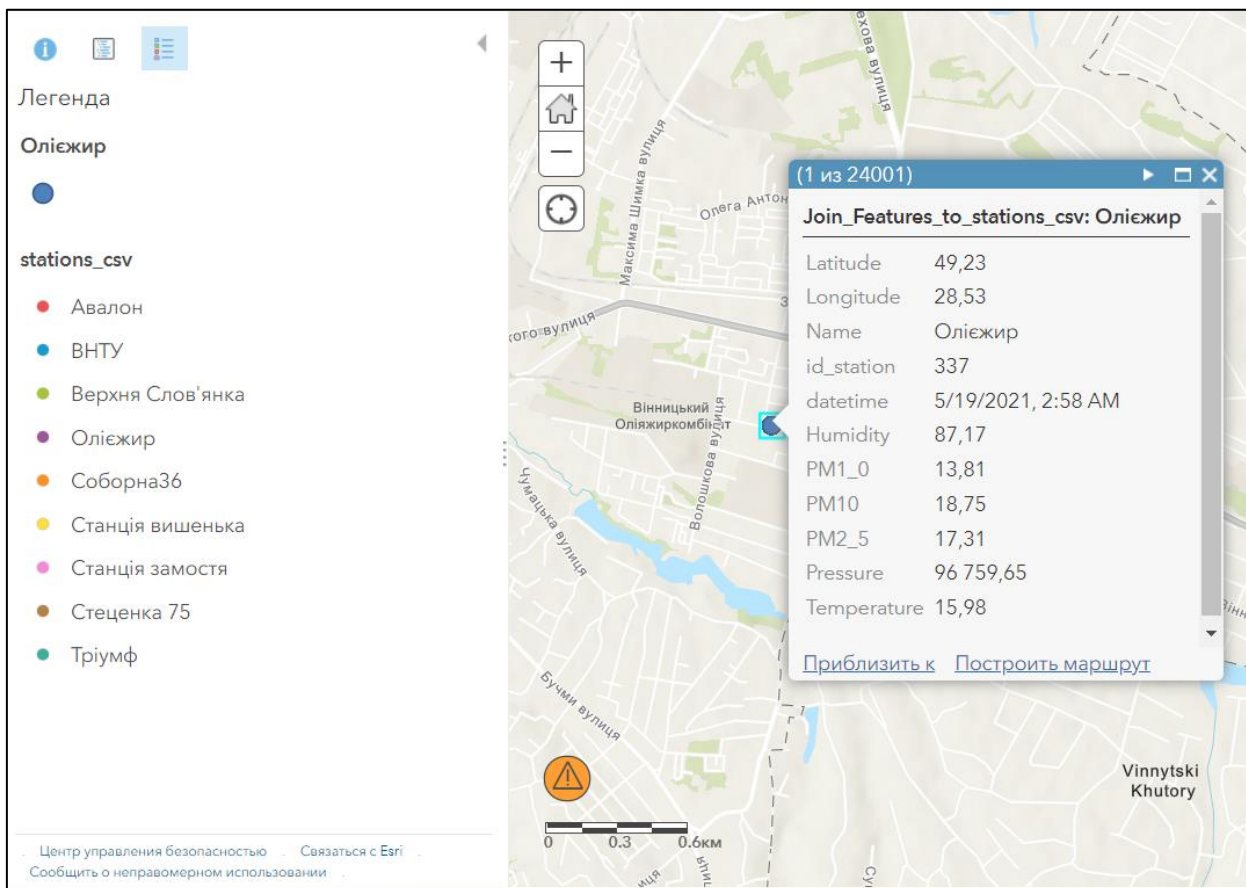


Рисунок 2.17 – Карта станції Олієжир

На рисунку бачимо координати даної точки, назву та номер (id) посту – Олієжир 337, дата та час коли було зафіксовано вимірювання показників, та нижче бачимо ознаки, за якими даний пост фіксував значення, а саме: humidity – вологість, pm1 – частинки пилу розміром 1 мкм, pm2,5 – частинки пилу розміром 2,5 мкм, pm10 – частинки пилу розміром 10 мкм, pressure – тиск та temperature – температура.

На рисунку 2.18 зображено карту станції Вишенька.

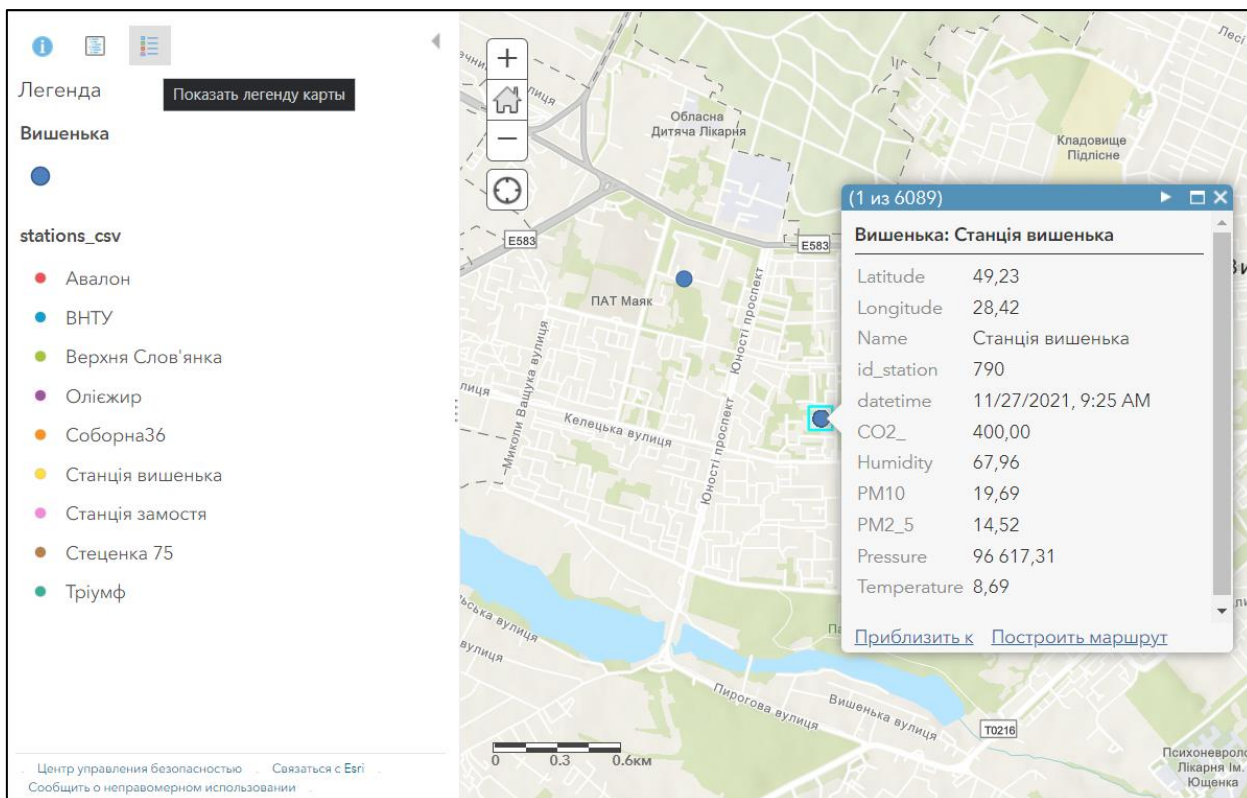


Рисунок 2.18 – Карта станції Вишенька

На рисунку бачимо координати даної точки, назву та номер (id) посту – Вишенька 790, дата та час коли було зафіксовано вимірювання показників, та нижче бачимо ознаки, за якими даний пост фіксував значення, а саме: humidity – вологість, CO2 – вуглекислий газ, pm2,5 – частинки пилу розміром 2,5 мкм, pm10 – частинки пилу розміром 10 мкм, pressure – тиск та temperature – температура.

На рисунку 2.19 зображено карту станції Стеценка 75.

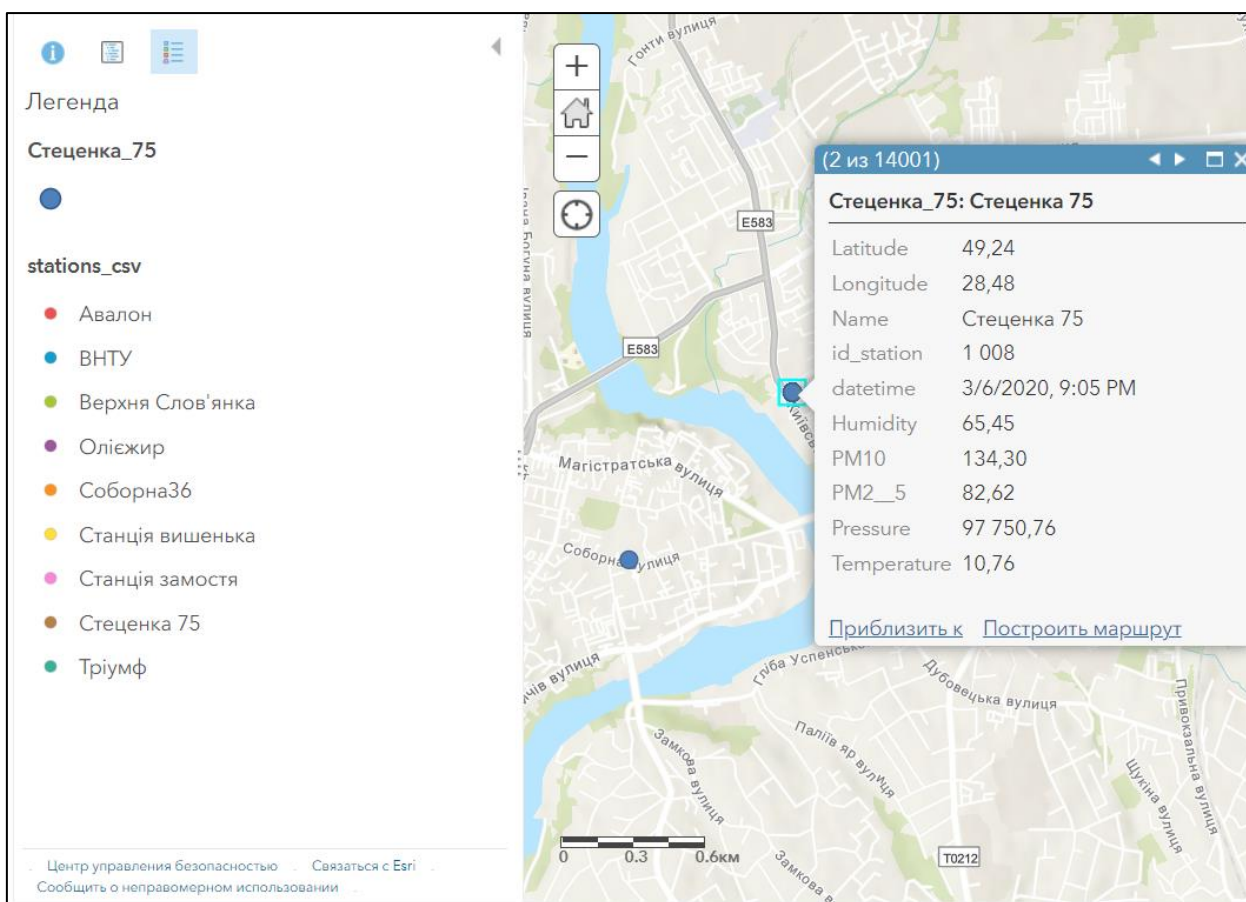


Рисунок 2.19 – Карта станції Стеценка 75

На рисунку бачимо координати даної точки, назву та номер (id) посту – Стеценка 75 1008, дата та час коли було зафіксовано вимірювання показників, та нижче бачимо ознаки, за якими даний пост фіксував значення, а саме: humidity – вологість, pm2,5 – частинки пилу розміром 2,5 мкм, pm10 – частинки пилу розміром 10 мкм, pressure – тиск та temperature – температура.

На рисунку 2.20 зображено карту станції Замостя.

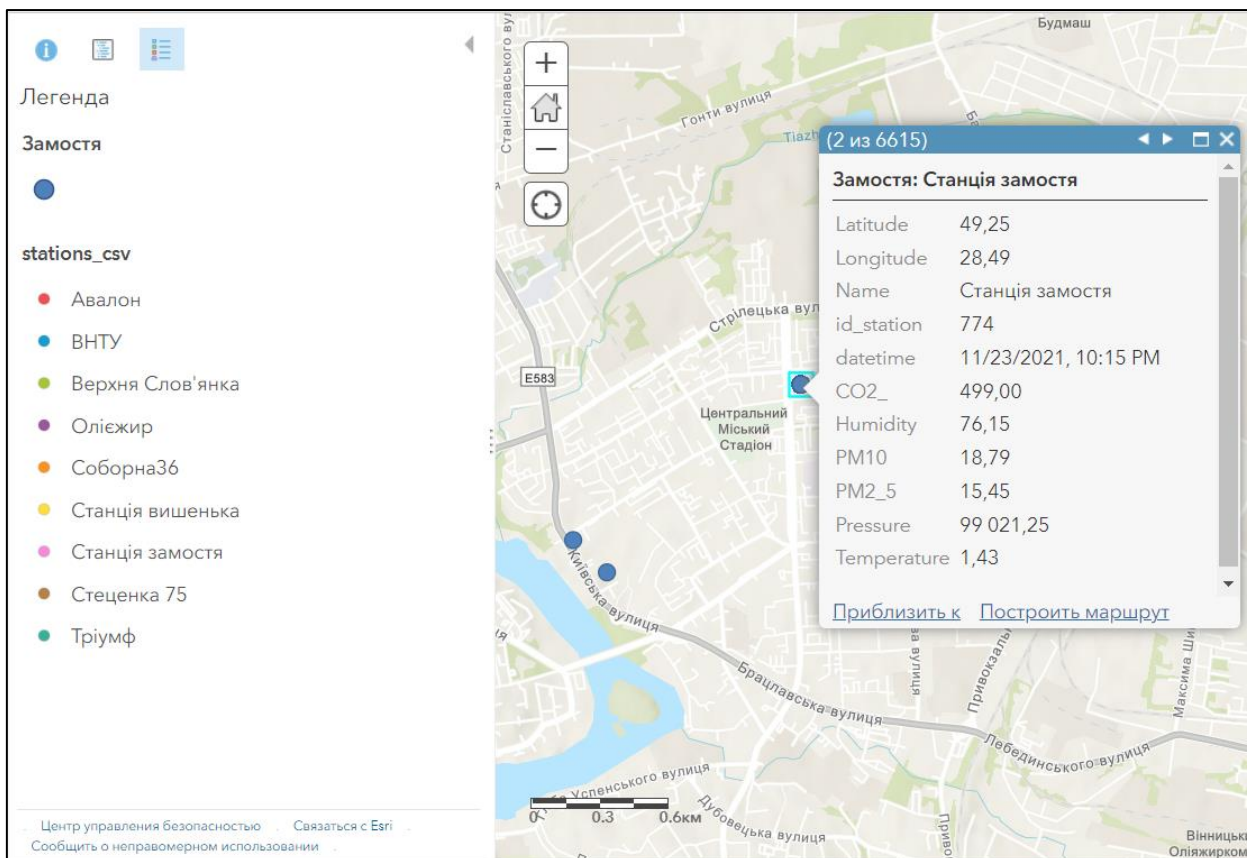


Рисунок 2.20 – Карта станції Замостя

На рисунку бачимо координати даної точки, назву та номер (id) посту – Замостя 774, дата та час коли було зафіксовано вимірювання показників, та нижче бачимо ознаки, за якими даний пост фіксував значення, а саме: CO2 – вуглекислий газ, humidity – вологість, pm2,5 – частинки пилу розміром 2,5 мкм, pm10 – частинки пилу розміром 10 мкм, pressure – тиск та temperature – температура.

На рисунку 2.21 зображено карту станції ВНТУ.

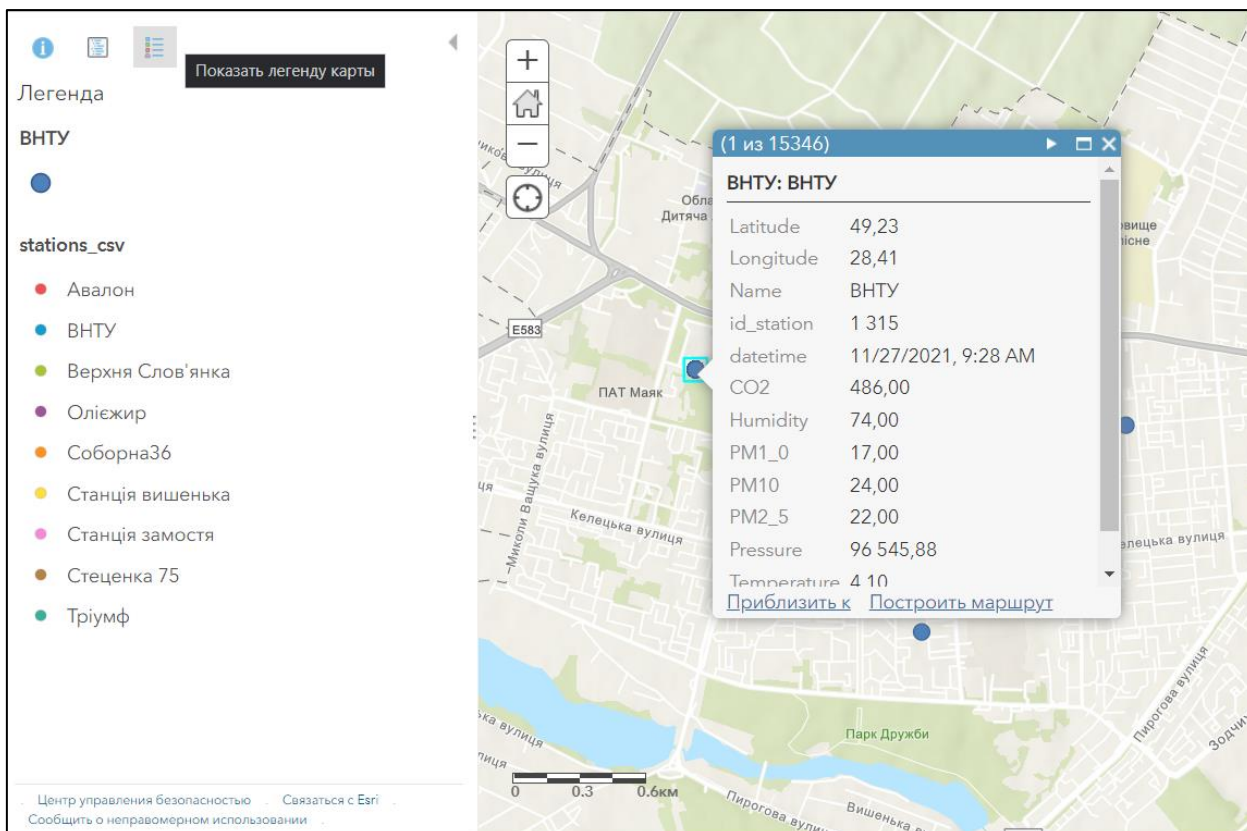


Рисунок 2.21 – Карта станції ВНТУ

На рисунку бачимо координати даної точки, назву та номер (id) посту – ВНТУ 1315, дата та час коли було зафіксовано вимірювання показників, та нижче бачимо ознаки, за якими даний пост фіксував значення, а саме: CO2 – вуглекислий газ, humidity – вологість, pm1 – частинки пилу розміром 1 мкм, pm2,5 – частинки пилу розміром 2,5 мкм, pm10 – частинки пилу розміром 10 мкм, pressure – тиск та temperature – температура.

На рисунку 2.22 зображено карту станції Авалон.

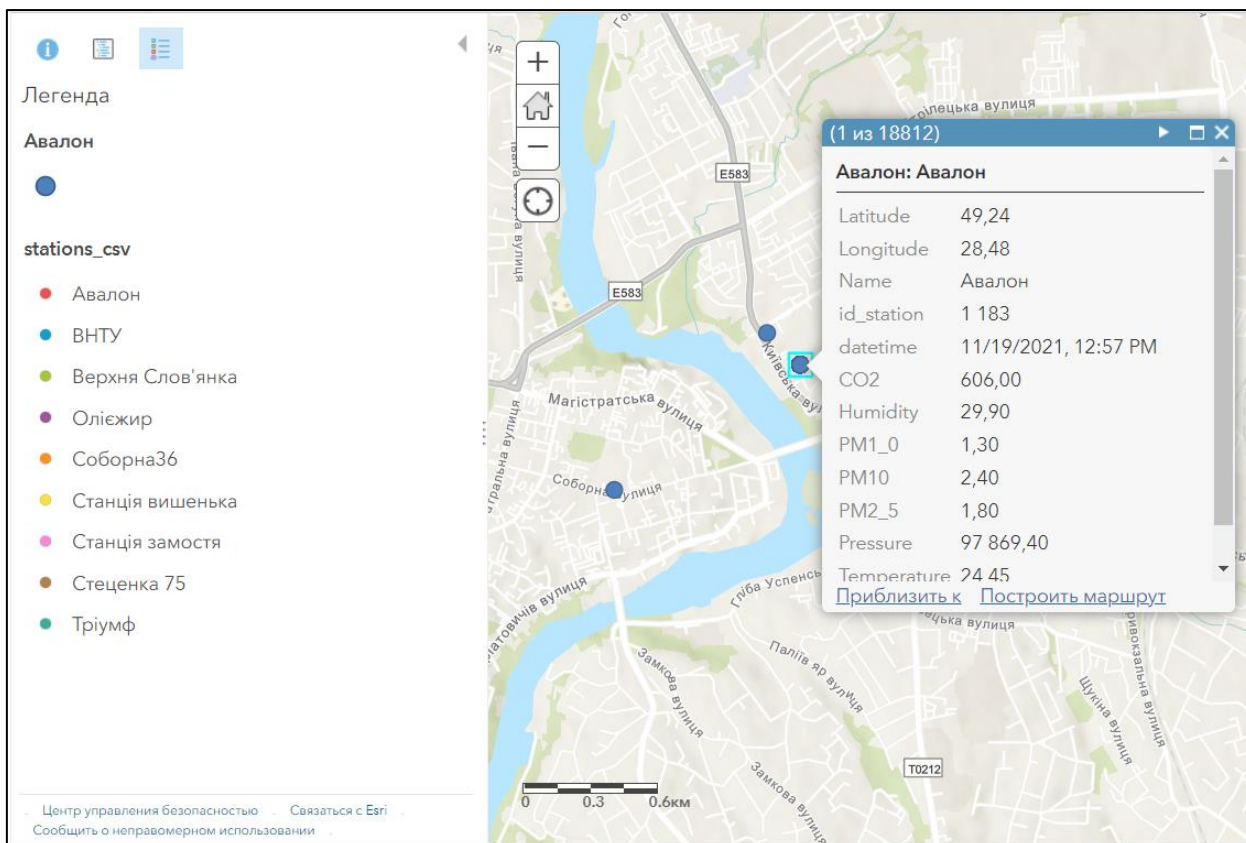


Рисунок 2.22 – Карта станції Авалон

На рисунку бачимо координати даної точки, назву та номер (id) посту – Авалон 1183, дата та час коли було зафіксовано вимірювання показників, та нижче бачимо ознаки, за якими даний пост фіксував значення, а саме: CO2 – вуглекислий газ, humidity – вологість, pm1 – частинки пилу розміром 1 мкм, pm2,5 – частинки пилу розміром 2,5 мкм, pm10 – частинки пилу розміром 10 мкм, pressure – тиск та temperature – температура.

На рисунку 2.23 зображено карту станції Тріумф.

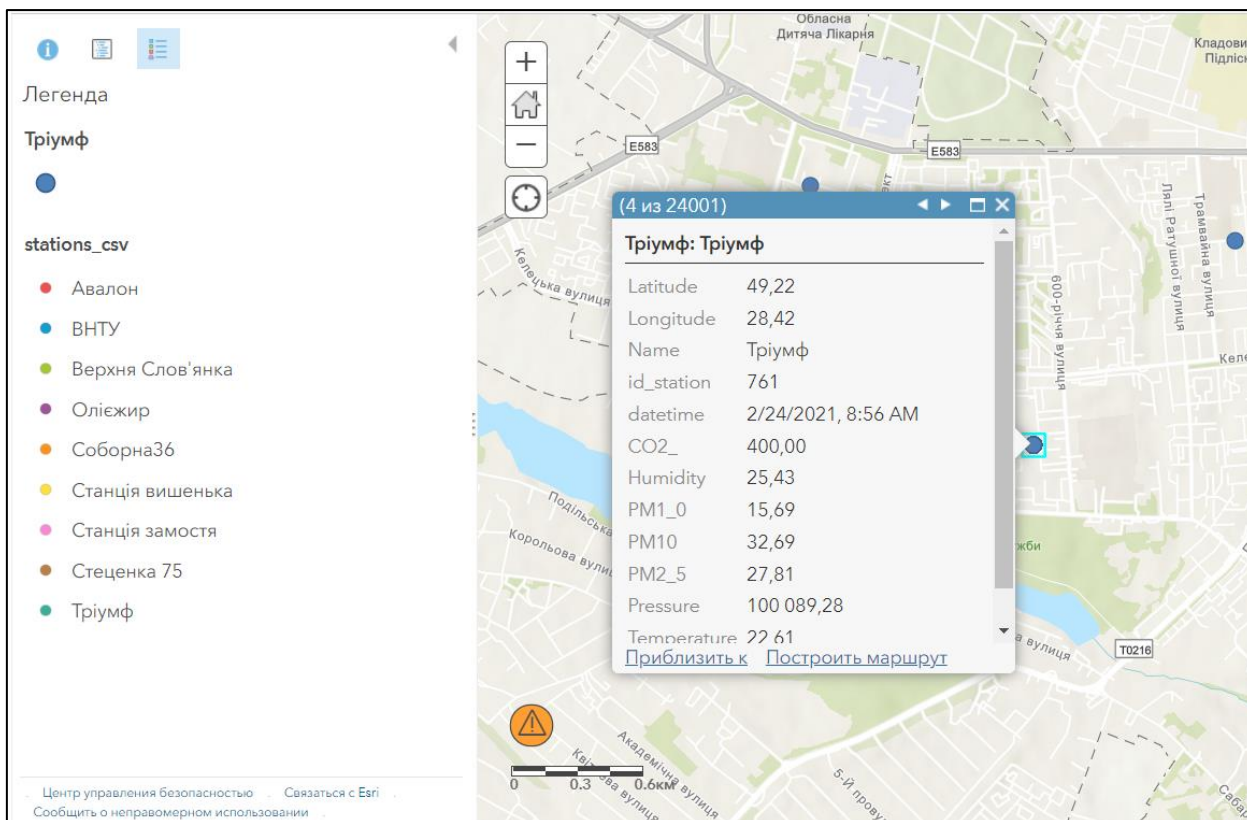


Рисунок 2.23 – Карта станції Тріумф

На рисунку бачимо координати даної точки, назву та номер (id) посту – Тріумф 761, дата та час коли було зафіксовано вимірювання показників, та нижче бачимо ознаки, за якими даний пост фіксував значення, а саме: CO2 – вуглекислий газ, humidity – вологість, pm1 – частинки пилу розміром 1 мкм, pm2,5 – частинки пилу розміром 2,5 мкм, pm10 – частинки пилу розміром 10 мкм, pressure – тиск та temperature – температура.

На рисунку 2.24 зображено карту станції Соборна 36.

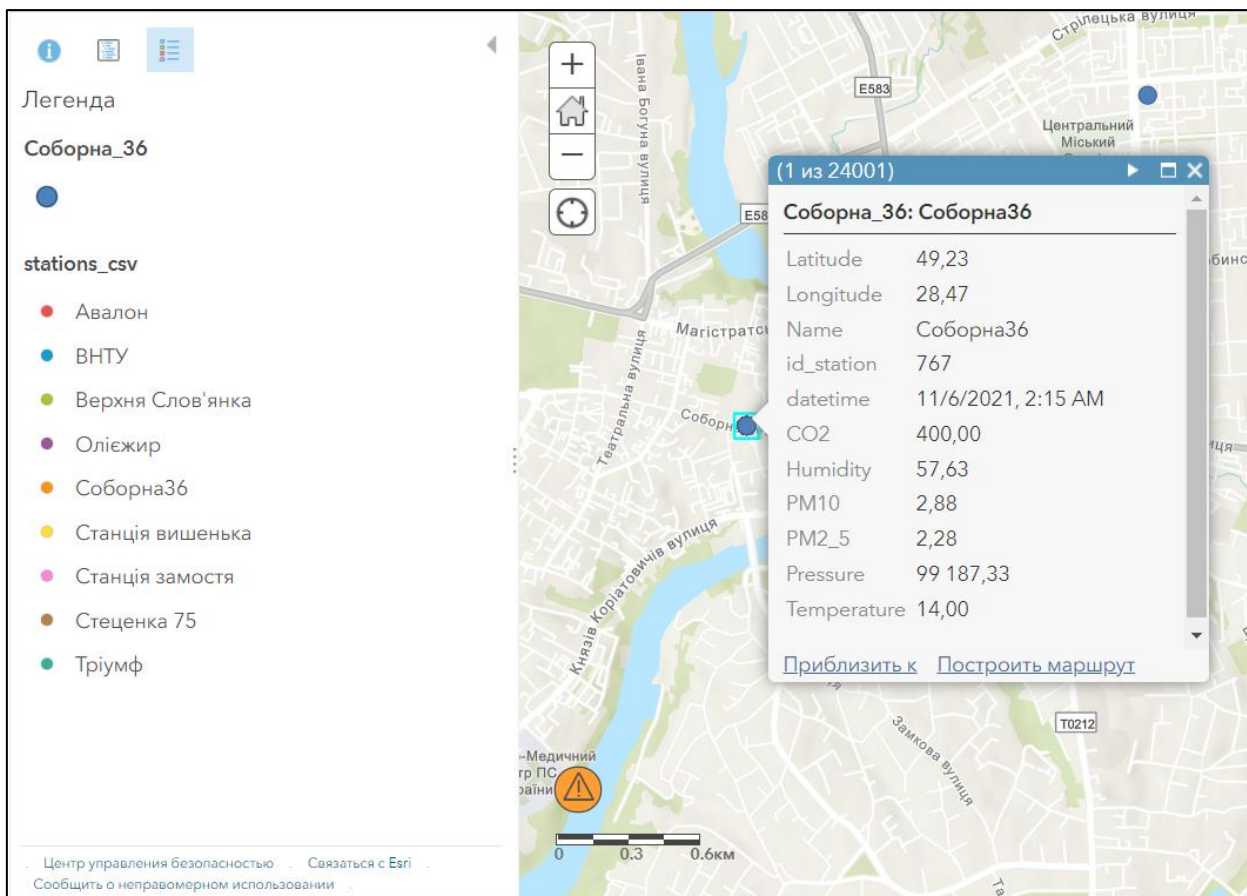


Рисунок 2.24 – Карта станції Соборна 36

На рисунку бачимо координати даної точки, назву та номер (id) посту – Соборна 36 767, дата та час коли було зафіксовано вимірювання показників, та нижче бачимо ознаки, за якими даний пост фіксував значення, а саме: CO2 – вуглекислий газ, humidity – вологість, pm2,5 – частинки пилу розміром 2,5 мкм, pm10 – частинки пилу розміром 10 мкм, pressure – тиск та temperature – температура.

На рисунку 2.25 зображено карту станції Верхня Солов'янка.

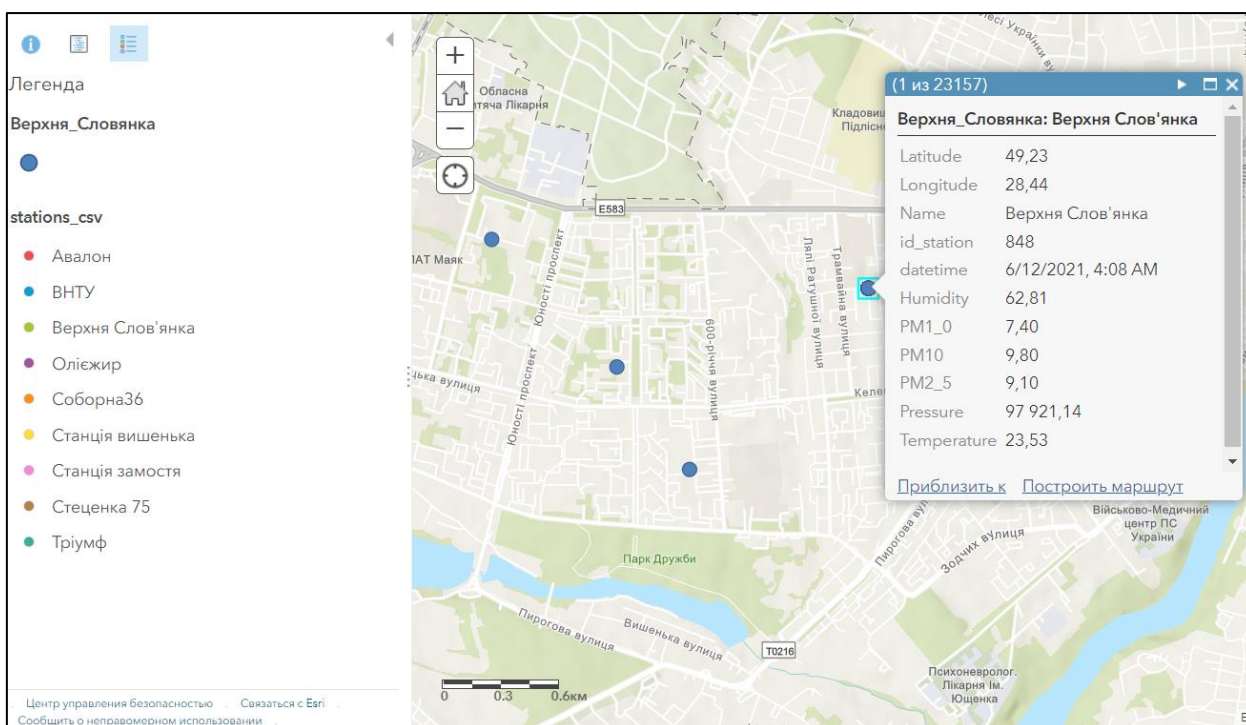


Рисунок 2.25 – Карта станції Верхня Солов'янка

На рисунку бачимо координати даної точки, назву та номер (id) посту – Верхня Солов'янка 848, дата та час коли було зафіксовано вимірювання показників, та нижче бачимо ознаки, за якими даний пост фіксував значення, а саме: humidity – вологість, pm1 – частинки пилу розміром 1 мкм, pm2,5 – частинки пилу розміром 2,5 мкм, pm10 – частинки пилу розміром 10 мкм, pressure – тиск та temperature – температура.

2.3 Висновки

У даному розділі було здійснено вибір оптимальних інформаційних технологій, огляд вхідних даних, відображення постів з набором даних на карті, та аналіз показників, що визначаються постами моніторингу якості атмосферного повітря.

3 ОБРОБКИ РЕЗУЛЬТАТІВ СПОСТЕРЕЖЕНЬ ТА ВІЗУАЛІЗАЦІЇ ДАНИХ У СЕРЕДОВИЩІ KAGGLE

3.1 Підготовка вхідних даних

У розділі 2.2 було детально описано станції та набір даних по кожній з цих станцій. Здійснимо геостатистичний аналіз та інтерполяцію у середовищі Kaggle [2]. Для початку, у середовище було завантажено набори даних по усіх станція у Вінницькій області. На рисунку 3.1 зображено завантажені набори дані з середовища SaveEcoBot.

	id_saveecobot	id_ecocity	network	locality	address	start_date	notes	lat	lng
0	14627	848.0	Eco-City	Vinnytsia city	2nd Konstantynovycha Ln, 4	14.12.2020	NaN	49.2307	28.4371
1	4152	767.0	Eco-City	Vinnytsia city	Soborna St, 36	26.06.2020	Vinnytsia City Concil	49.2329	28.4704
2	1008	NaN	SaveDnipro	Vinnytsia city	Stetsenka St, 75	25.05.2019	NaN	49.2393	28.4799
3	17796	NaN	luftdaten.info	Vinnytsia city	Kyiv's'ka St, 5	16.05.2021	NaN	49.2380	28.4820
4	15952	337.0	Eco-City	Vinnytsia city	3rd Chekhova Passage, 21	23.01.2021	NaN	49.2268	28.5256

Рисунок 3.1 – Завантаження даних середовища SaveEcoBot

Бачимо, що з даного середовища було завантажено 5 таблиць з даними, та усі з міста Вінниця. Здійснивши аналіз вияснили, що кількість даних для станції 848 становить 4600, дані для станції 767 – не визначені, кількість даних для станції 1008 становить 6717, кількість даних для станції 17796 становить 1968, кількість даних для станції 337 – 4791. Також, аналіз показав, що у наборі немає пустих комірок, тобто усі дані визначені, пропущених немає.

На рисунку 3.2 зображено завантажені станції з середовища Eco City.

	id_saveecobot	id_ecocity	network	locality	address	start_date	lat	lng	notes	source
0	15952.0	337	Eco-City	Vinnytsia city	3rd Chekhova Passage, 21	23.01.2021	49.2267	28.5259	Oil and fat factory	https://zoom=
1	4238.0	650	Eco-City	Vinnytsia region	Turbiv town	02.08.2020	49.3435	28.7298	NaN	https://zoom=
2	4152.0	767	Eco-City	Vinnytsia city	Soborna St, 36	26.06.2020	49.2329	28.4704	Vinnytsia City Concil	https://zoom=
3	NaN	774	Eco-City	Vinnytsia city	Vinnytsia, Zamostyanska Street, 7/35	NaN	49.2455	28.4936	Transparent office	https://zoom=
4	NaN	790	Eco-City	Vinnytsia city	Vinnytsia, Cosmonauts Avenue, 30	NaN	49.2268	28.4194	Transparent office	https://zoom=
5	14627.0	848	Eco-City	Vinnytsia city	2nd Konstanynovycha Ln, 4	14.12.2020	49.2307	28.4371	NaN	https://zoom=
6	14855.0	1172	Eco-City	Vinnytsia region	Ivaniv village	28.12.2020	49.4838	28.3636	NaN	https://zoom=
7	NaN	1315	Eco-City	Vinnytsia city	Vinnytsia, str. Khmelnytsky shose, 95, VNTU, b...	16.11.2021	49.2330	28.4100	Vinnytsia National Technical University	https://zoom=
8	NaN	1183	Eco-City	Vinnytsia city	Vinnytsia, str. Tram, 3	NaN	49.2311	28.4371	Avalon	https://
9	NaN	761	Eco-City	Vinnytsia city	Vinnytsia, 600th anniversary, 66V	NaN	49.2222	28.4242	Triumph	https://

Рисунок 3.2 – Завантаження даних середовища Eco City

Рисунок 3.2 аналогічний до попереднього, з середовища Eco City було додано 10 таблиць. Аналіз показав, що у даних таблицях пропущених значень немає, усі комірки заповнені.

Оскільки у завданні потрібно виконати інтерполяцію для міста Вінниця, дані було відсортовано, аби не враховувати всю Вінницьку область. На рисунку 3.3 зображена таблиця постів у місті Вінниця.

	id_station	value	lat	lng	network
0	EcoCity_1183	1162.413898	49.2311	28.4371	Eco-City
1	EcoCity_1315	57.947368	49.2330	28.4100	Eco-City
2	EcoCity_337	402.385263	49.2267	28.5259	Eco-City
3	EcoCity_761	584.883654	49.2222	28.4242	Eco-City
4	EcoCity_767	131.162373	49.2329	28.4704	Eco-City
5	EcoCity_774	25.046724	49.2455	28.4936	Eco-City
6	EcoCity_790	31.416316	49.2268	28.4194	Eco-City
7	EcoCity_848	114.972500	49.2307	28.4371	Eco-City
8	SaveEcoBot_1008	374.180000	49.2393	28.4799	SaveDnipro
9	SaveEcoBot_17796	187.460000	49.2380	28.4820	luftdaten.info

Рисунок 3.3 – Таблиця постів у місті Вінниця

Можемо спостерігати, що у таблиці відображаються 10 постів, проте у попередніх розділах вказувалось, що постів лише 9. Пост під номером 17796 має такі ж координати як 1183 Авалон, тобто він переїхав за іншою адресою, однак актуальні дані у таблиці 1183.

На рисунку 3.4 відображена програмна реалізація відбору даних для інтерполяції.

```
# Selection data for interpolation
if type_agg == 'mean':
    data = df[df['ds']==datetime.datetime.fromisoformat(datetime_analysis)].reset_index(drop=True)
else:
    # type_agg == 'max':
    datetime_analysis = 'all time'
    data = df[['value', 'id_station']].groupby(by=['id_station']).max()
    data = data.reset_index(drop=False)
    data = pd.merge(data, df[['id_station', 'lat', 'lng', 'network']], how = 'left', on = 'id_station').drop_duplicates().reset_index(drop=True)

x = data.lng.values
y = data.lat.values
z = data.value.values
fig = plt.figure()
plt.scatter(x, y)
plt.title(f'Stations in Vinnytsia region with data for {indicator_name} in {datetime_analysis}')
display(data)
plt.show()
```

Рисунок 3.4 – Програмна реалізація відбору даних для інтерполяції

На рисунку 3.5 зображено графік розташування постів за координатами один відносно одного.

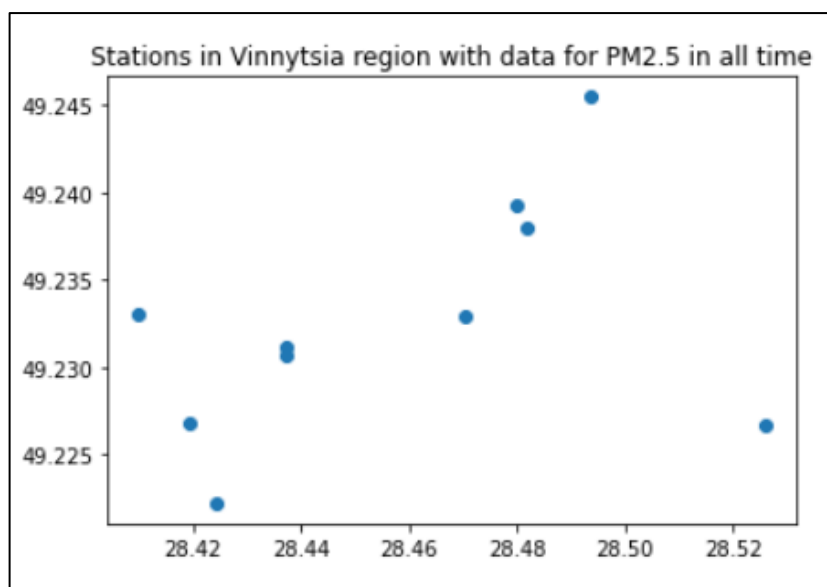


Рисунок 3.5 – Графік розташування постів за координатами

По осі ординат взято значення *latitude* – широта, а по осі абсцис *longitude* – довжина. Аби з'ясувати, який пост відповідає точці на даному рисунку, можливо скористатись рисунком 3.3, підставивши значення *lat* та *lng*.

3.2 Геостатистичний аналіз за допомогою SciKit-GStat

Для початку потрібно відсортувати стовпці з таблиці 3.3, залишити лише дату та час (*ds*), значення показників (*value*), мережа (*network*), номер станції (*d_station*), ширину (*lat*) та довжину (*lng*). На рисунку 3.6 зображена таблиця даних, підготовлених для геостатистичного аналізу.

	<i>ds</i>	<i>value</i>	<i>network</i>	<i>id_station</i>	<i>lat</i>	<i>lng</i>
0	2021-11-25 01:00:00	39.313000	Eco-City	EcoCity_337	49.2267	28.5259
1	2021-11-25 01:00:00	12.699000	Eco-City	EcoCity_767	49.2329	28.4704
2	2021-11-25 01:00:00	15.428000	Eco-City	EcoCity_774	49.2455	28.4936
3	2021-11-25 01:00:00	18.262000	Eco-City	EcoCity_790	49.2268	28.4194
4	2021-11-25 01:00:00	30.206364	Eco-City	EcoCity_848	49.2307	28.4371
5	2021-11-25 01:00:00	28.200000	Eco-City	EcoCity_1315	49.2330	28.4100
6	2021-11-25 01:00:00	6.905000	Eco-City	EcoCity_761	49.2222	28.4242

Рисунок 3.6 – Підготовлені дані для геостатистичного аналізу

Далі на основі таблиці 3.6 було побудовано графік розташування точок на основі значень по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 25.11.2021 о 00:00:00 (рис. 3.7).

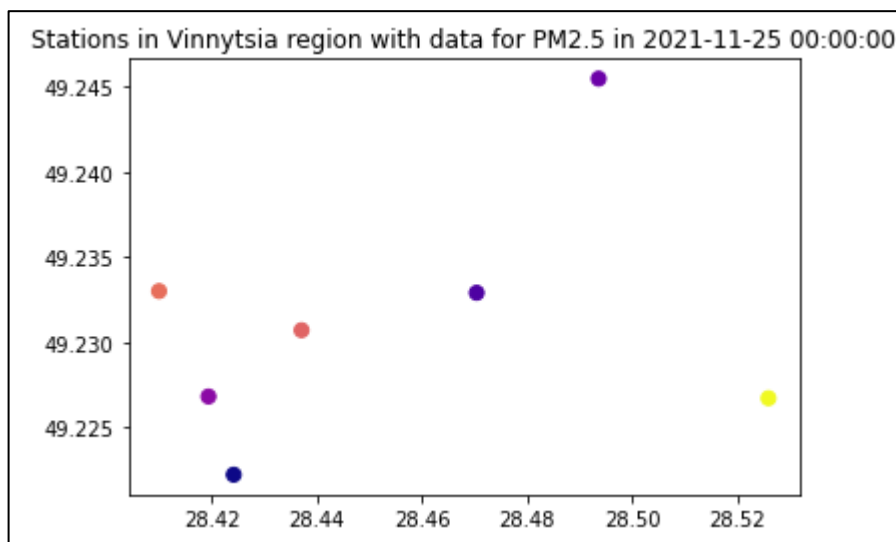


Рисунок 3.7 – графік розташування точок на основі значень по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 25.11.2021 о 00:00:00

На рисунку 3.8 зображена сферична варіограма по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 25.11.2021 о 00:00:00.

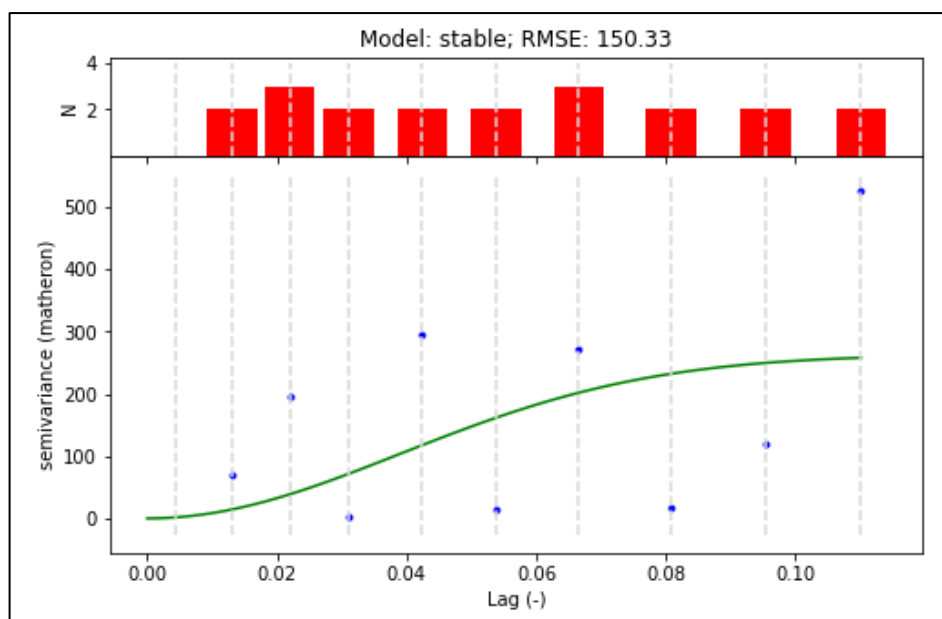


Рисунок 3.8 – Сферична варіограма по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 25.11.2021 о 00:00:00

На рисунку 3.8 можливо побачити на скільки далеко чи близько один від одного знаходяться пости відносно широти.

Було побудовано графік розташування точок на основі значень по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 25.11.2021 о 01:00:00 (рис. 3.9).

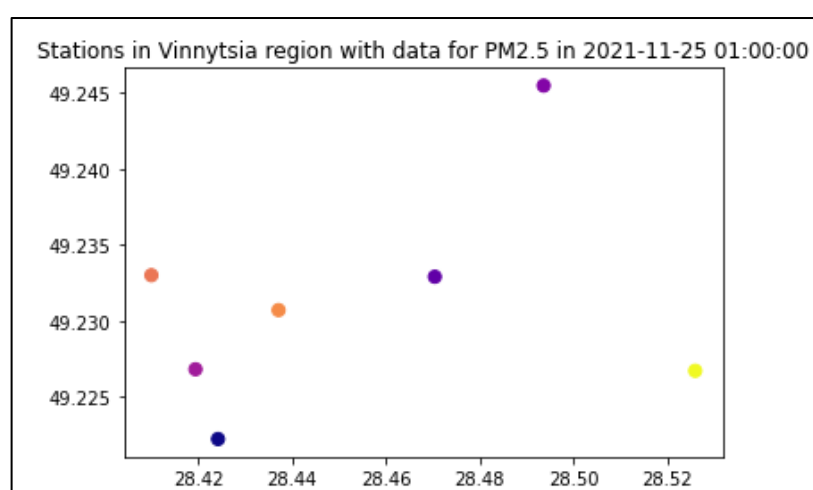


Рисунок 3.9 – графік розташування точок на основі значень по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 25.11.2021 о 01:00:00

На рисунку 3.10 зображена сферична варіограма по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 25.11.2021 о 01:00:00.

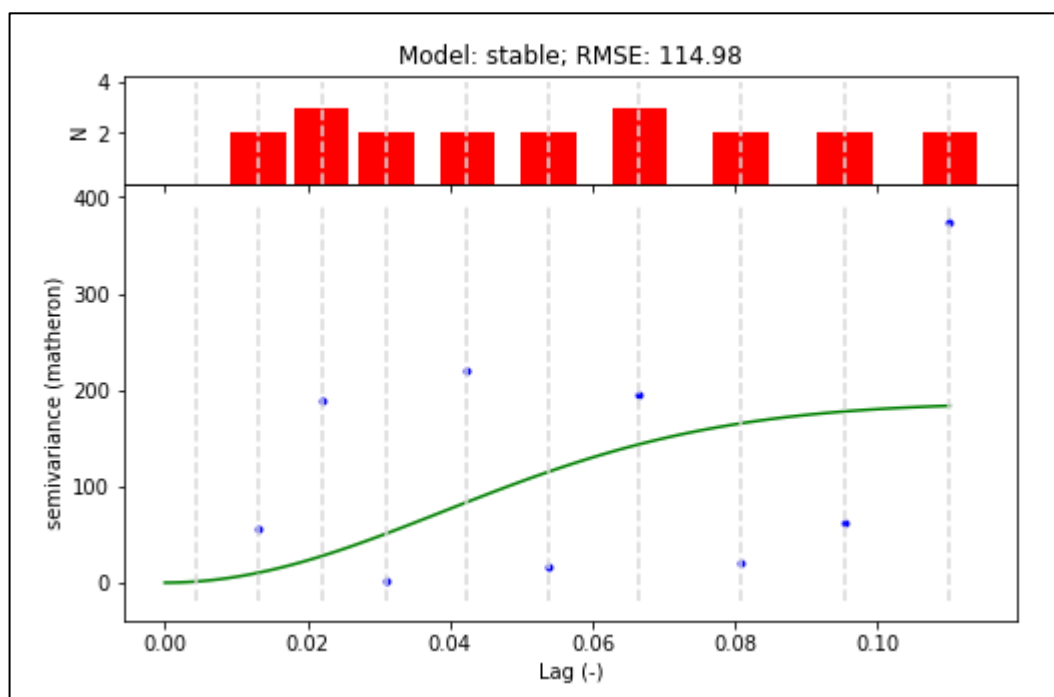


Рисунок 3.10 – Сферична варіограма по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 25.11.2021 о 01:00:00

Було побудовано графік розташування точок на основі значень по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 25.11.2021 о 06:00:00 (рис. 3.11).

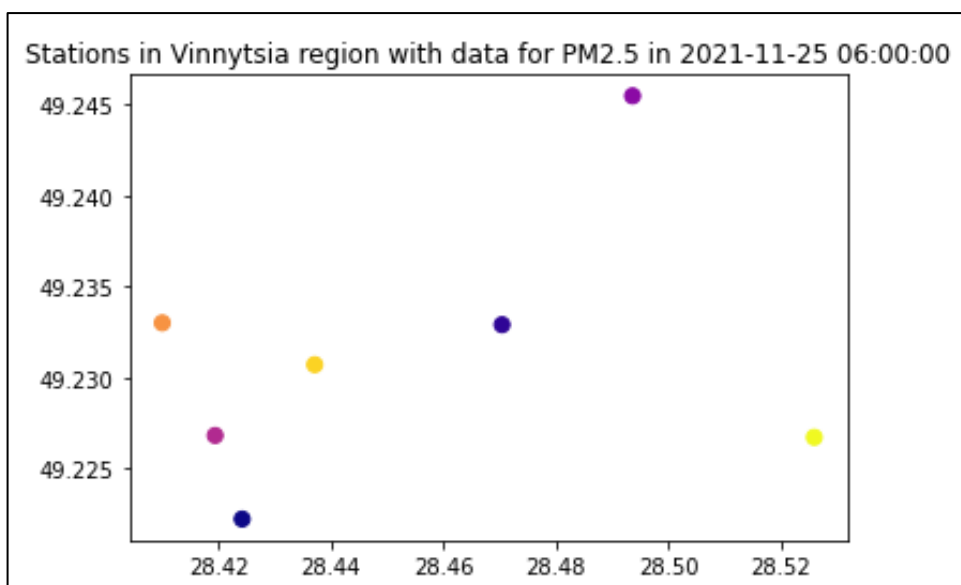


Рисунок 3.11 – графік розташування точок на основі значень по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 25.11.2021 о 06:00:00

На рисунку 3.12 зображена сферична варіограма по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 25.11.2021 о 06:00:00.

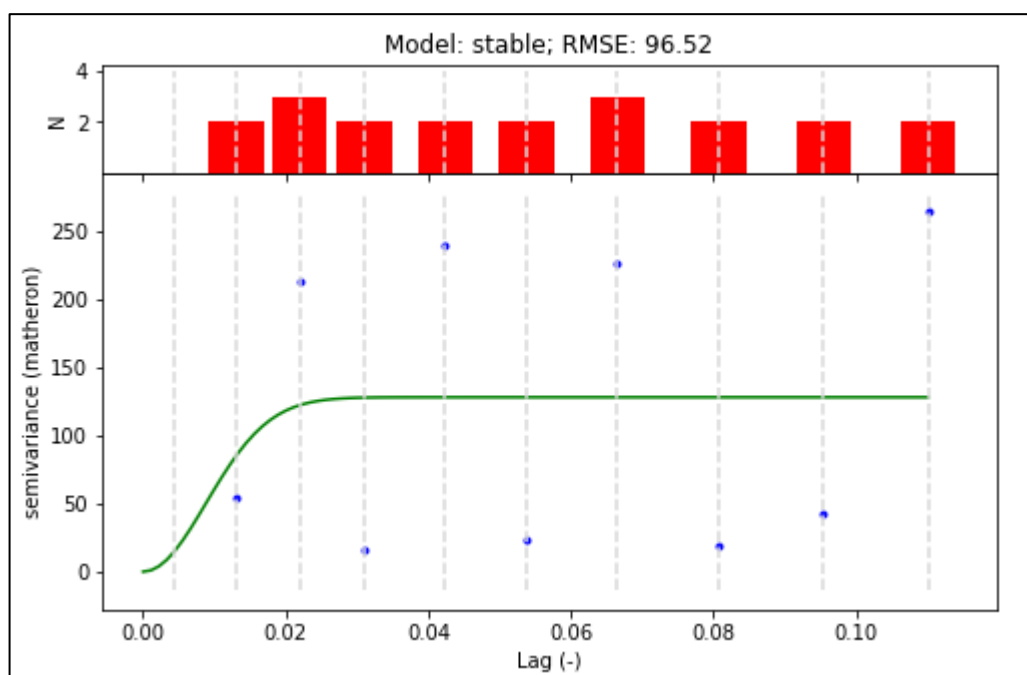


Рисунок 3.12 – Сферична варіограма по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 25.11.2021 о 06:00:00

Було побудовано графік розташування точок на основі значень по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 25.11.2021 о 07:00:00 (рис. 3.13).

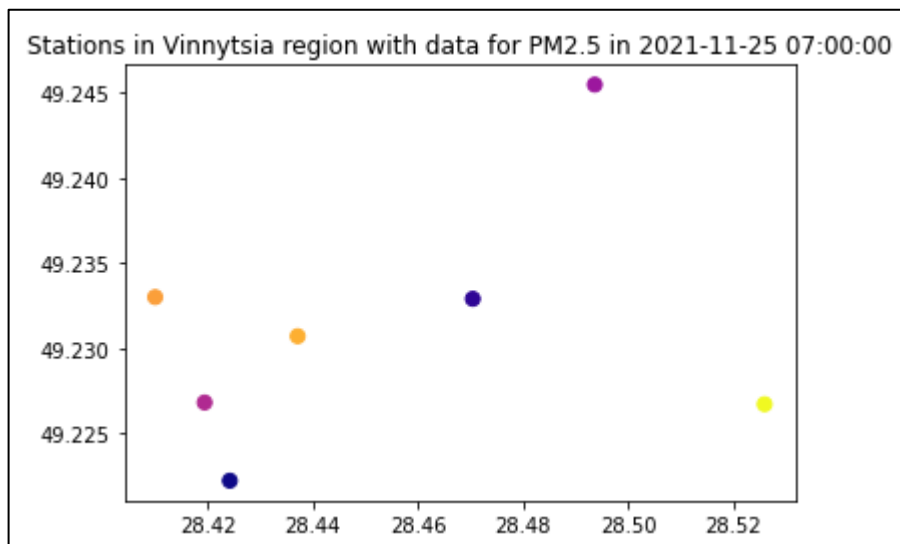


Рисунок 3.13 – графік розташування точок на основі значень по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 25.11.2021 о 07:00:00

На рисунку 3.14 зображена сферична варіограма по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 25.11.2021 о 07:00:00.

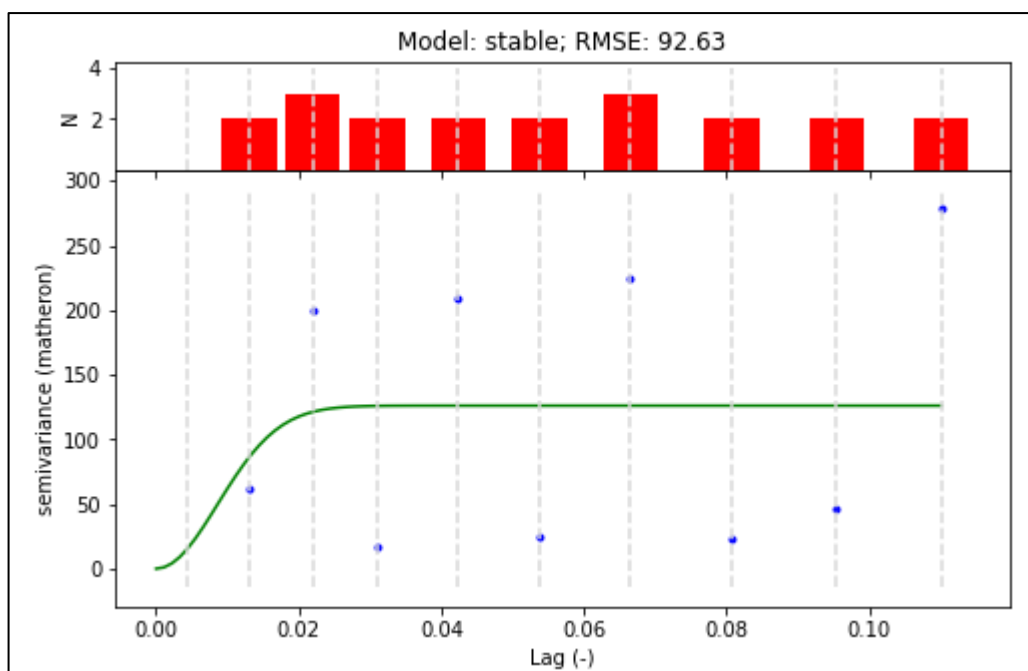


Рисунок 3.14 – Сферична варіограма по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 25.11.2021 о 07:00:00

Було побудовано графік розташування точок на основі значень по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 25.11.2021 о 08:00:00 (рис. 3.15).

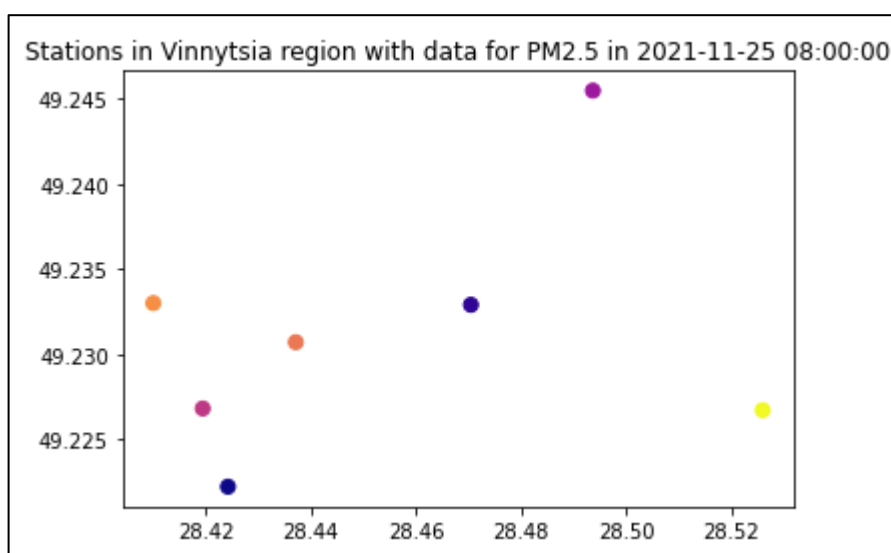


Рисунок 3.15 – графік розташування точок на основі значень по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 25.11.2021 о 08:00:00

На рисунку 3.16 зображена сферична варіограма по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 25.11.2021 о 08:00:00.

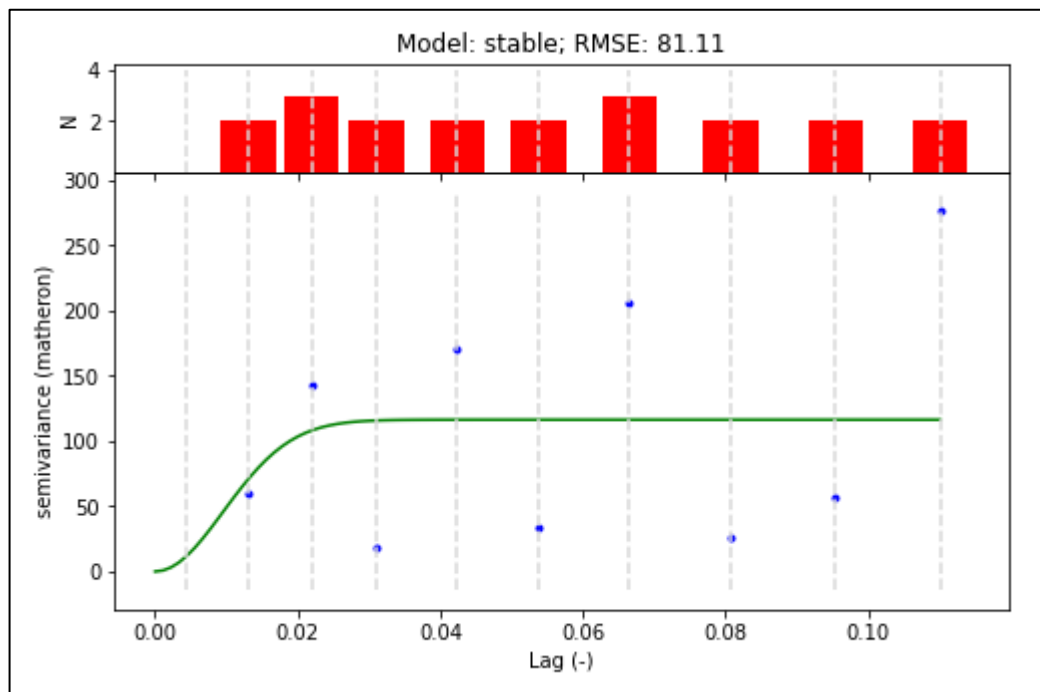


Рисунок 3.16 – Сферична варіограма по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 25.11.2021 о 08:00:00

Було побудовано графік розташування точок на основі значень по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 25.11.2021 о 09:00:00 (рис. 3.17).

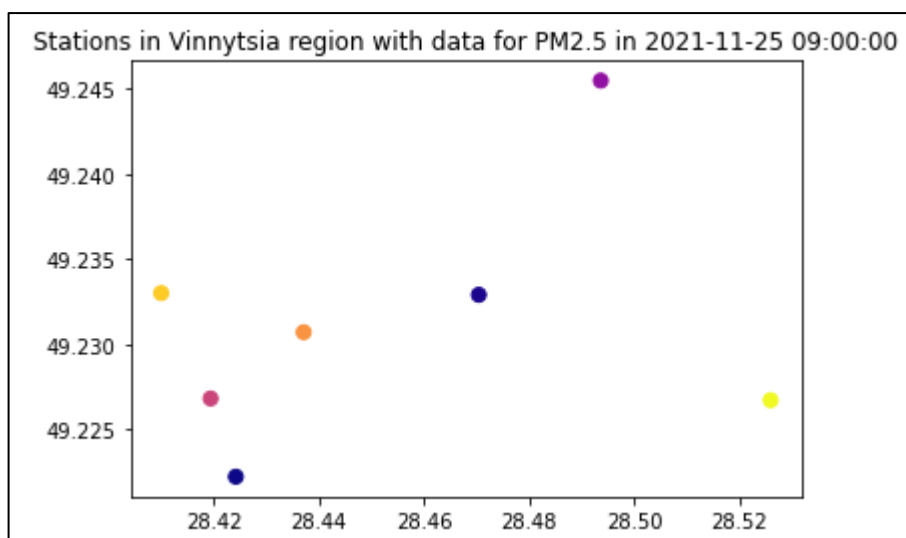


Рисунок 3.17 – графік розташування точок на основі значень по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 25.11.2021 о 09:00:00

На рисунку 3.18 зображена сферична варіограма по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 25.11.2021 о 09:00:00.

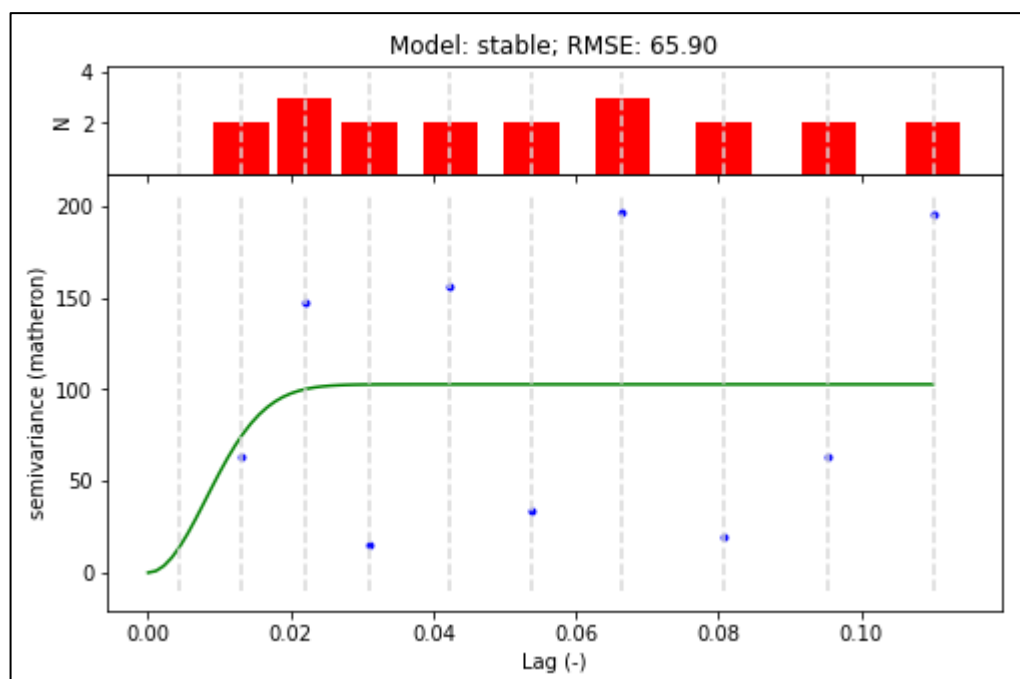


Рисунок 3.18 – Сферична варіограма по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 25.11.2021 о 09:00:00

Було побудовано графік розташування точок на основі значень по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 25.11.2021 о 10:00:00 (рис. 3.19).

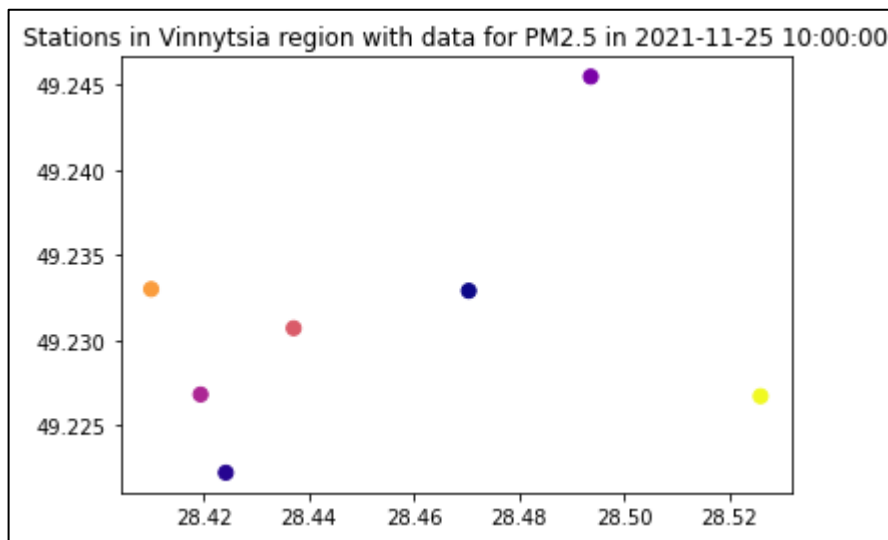


Рисунок 3.19 – графік розташування точок на основі значень по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 25.11.2021 о 10:00:00

На рисунку 3.20 зображена сферична варіограма по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 25.11.2021 о 10:00:00.

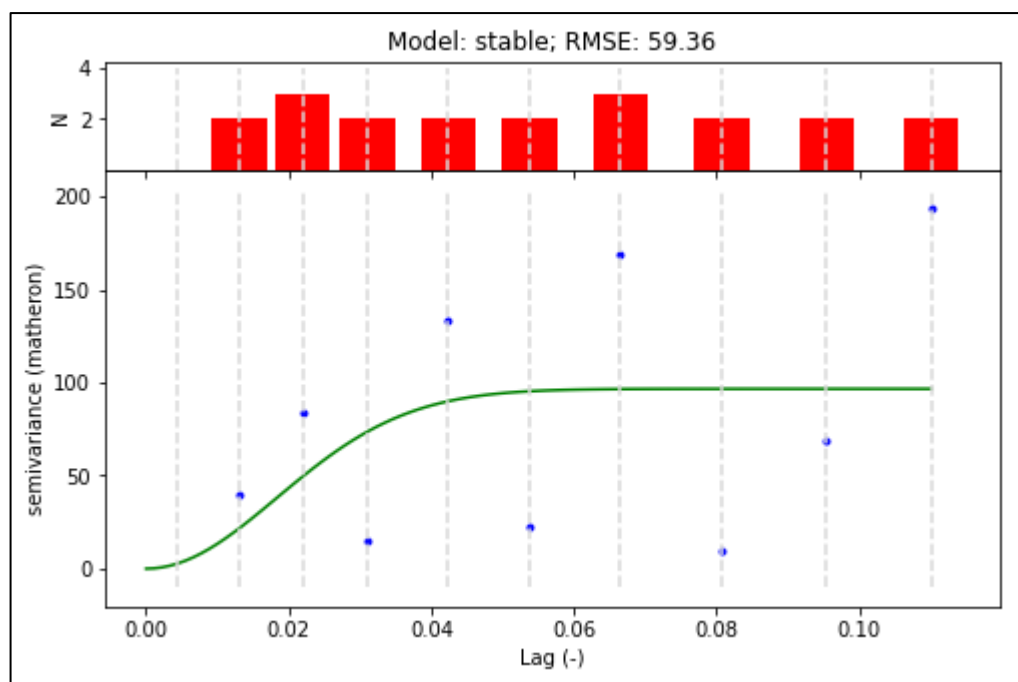


Рисунок 3.20 – Сферична варіограма по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 25.11.2021 о 10:00:00

Було побудовано графік розташування точок на основі значень по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 25.11.2021 о 11:00:00 (рис. 3.21).

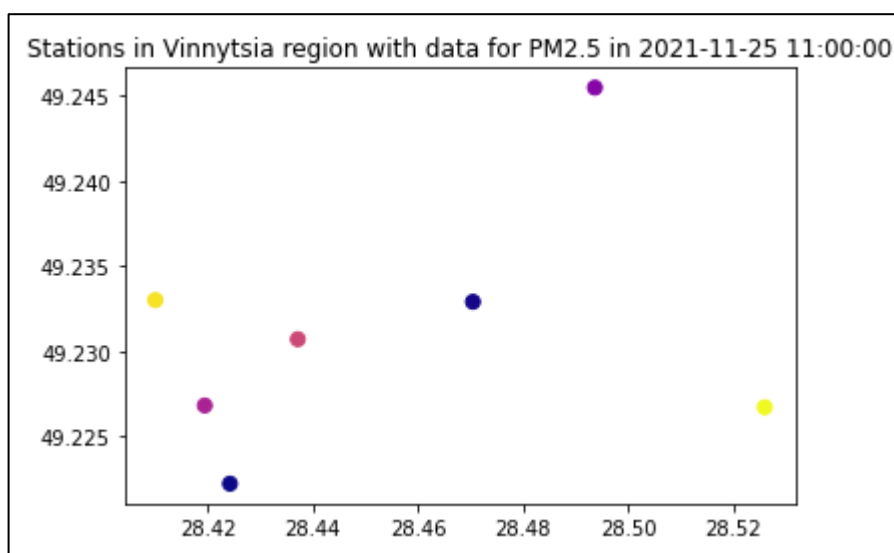


Рисунок 3.21 – графік розташування точок на основі значень по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 25.11.2021 о 11:00:00

На рисунку 3.22 зображена сферична варіограма по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 25.11.2021 о 11:00:00.

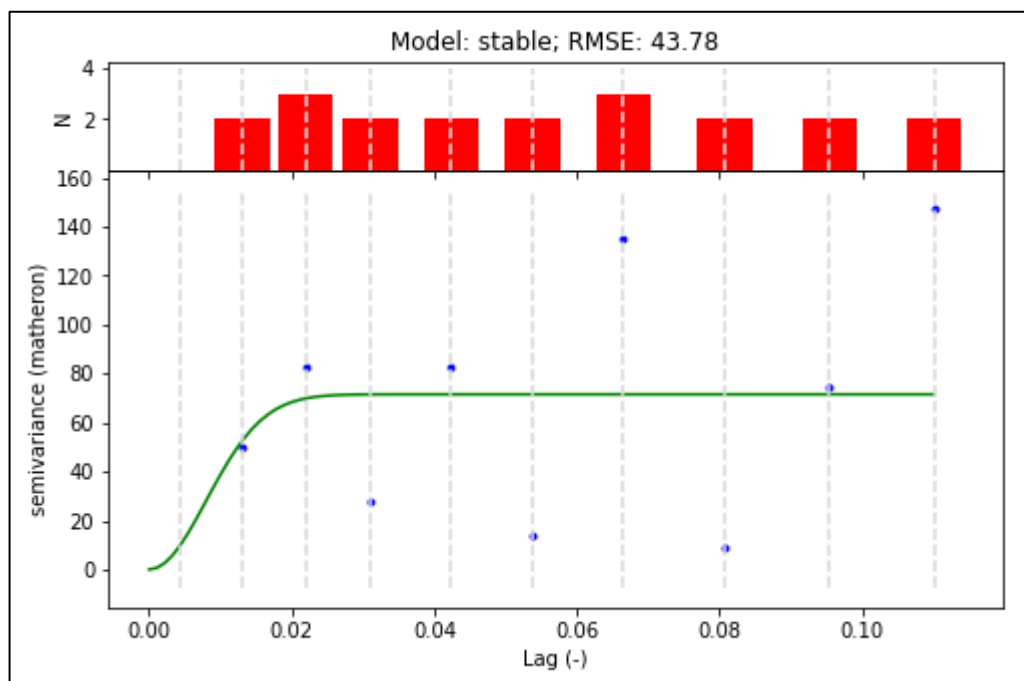


Рисунок 3.22 – Сферична варіограма по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 25.11.2021 о 11:00:00

Було побудовано графік розташування точок на основі значень по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 25.11.2021 о 12:00:00 (рис. 3.23).

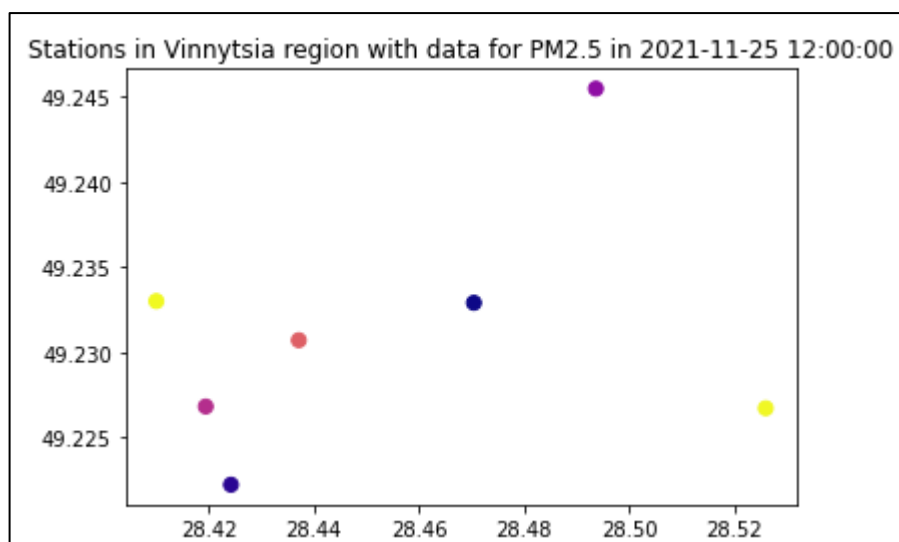


Рисунок 3.23 – графік розташування точок на основі значень по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 25.11.2021 о 12:00:00

На рисунку 3.24 зображена сферична варіограма по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 25.11.2021 о 12:00:00.

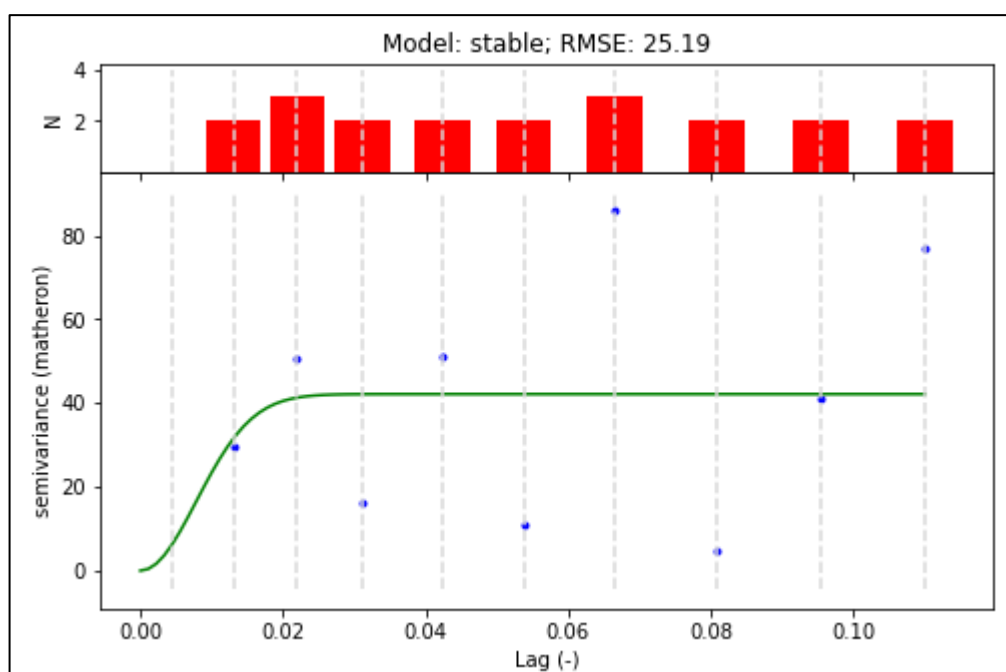


Рисунок 3.24 – Сферична варіограма по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 25.11.2021 о 12:00:00

Варіограми будуються з великою похибкою, оскільки є значна дисперсія розкиду між значеннями. Однією з причин може бути те, що, одночасно, і дисперсія (розкид значень) – завелика, і точок – замало. Тому можна зробити висновок про те, що треба ставити по-більше постів там, де є суттєві зміни якості атмосферного повітря, а по-друге, там – де постів замало (на півдні і на півночі міста).

3.3 Лінійна інтерполяція за допомогою `scipy.interpolate.interp2d`

На рисунку 3.25 зображена програмна реалізація лінійної візуалізації.

```
# Visualization
fig = plt.figure(figsize=(12,10))
plt.contourf(X, Y, Z)

plt.scatter(xseb, yseb, c='gray', s=100, label='SaveEcoBot')
for i in range(len(xseb)):
    plt.annotate(" "+numseb[i], xy=(xseb[i], yseb[i]), textcoords='data')

plt.scatter(xeco, yeco, c='k', s=100, label='EcoCity')
plt.colorbar()
for i in range(len(xeco)):
    plt.annotate(" "+numeco[i], xy=(xeco[i], yeco[i]), textcoords='data')

plt.axis()
type_agg_str = 'average' if type_agg=='mean' else 'maximum'
time_agg_str = 'hour' if time_interval=='H' else 'D'
plt.title(f'Stations in Vinnytsia region with {type_agg_str} data per {time_agg_str} for {indicator_name} in {datetime_analysis} (maximum value = {round(data.value.max(),2)})')
plt.legend(loc='best')
plt.grid()
plt.show()
```

Рисунок 3.25 – Програмна реалізація лінійної візуалізації

Було побудовано графіки лінійної інтерполяції для станцій у місті Вінниця за показником $PM_{2.5}$ мкм та максимальних значень за весь період (рис. 3.26).

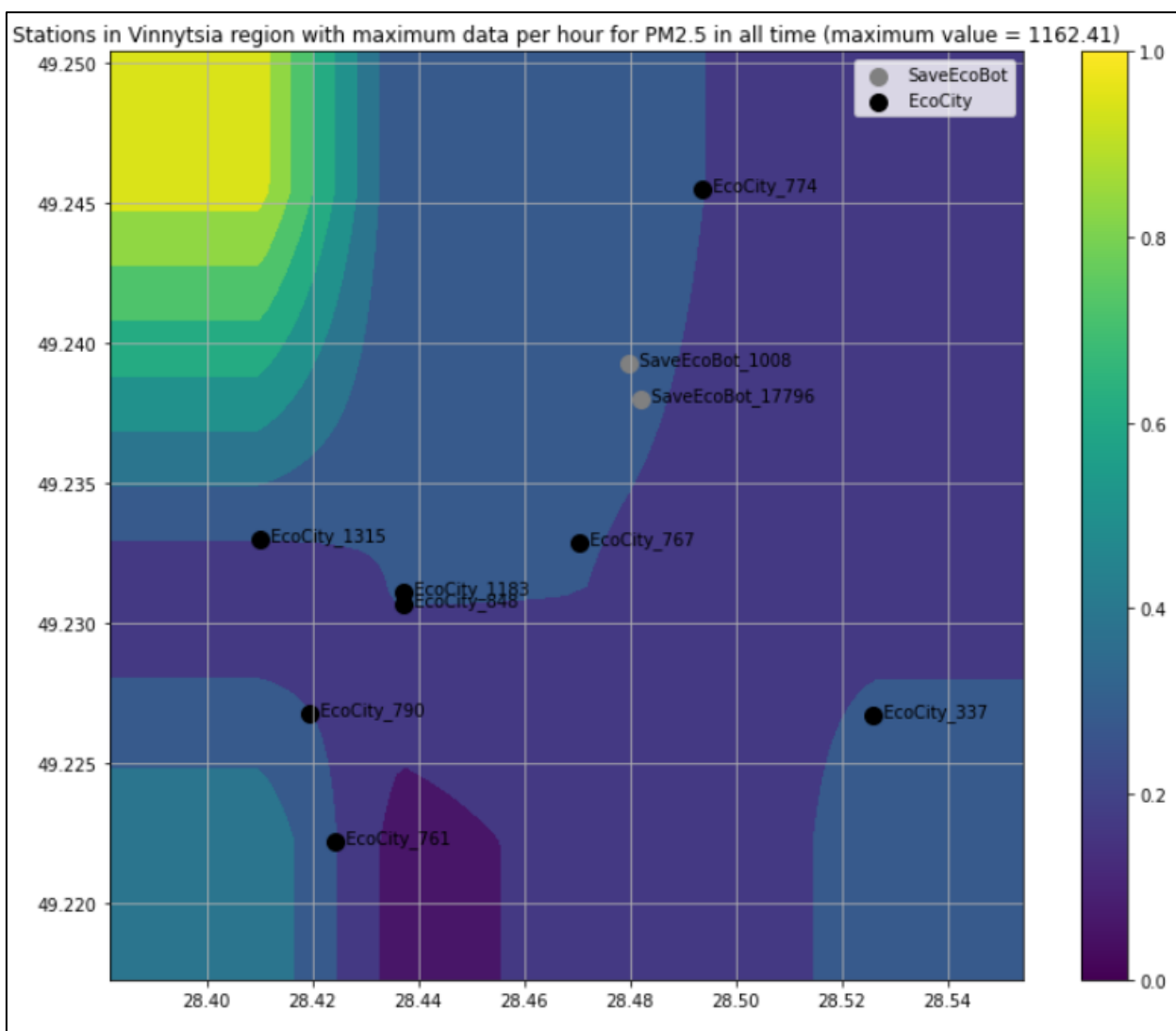


Рисунок 3.26 – Графік лінійної інтерполяції

На рисунку чорним кольором розмічені пости моніторингу Eco City, а сірим – SaveEcoBot.

3.4 Висновки

На основі зібраних даних спостережень з постів мережі громадського моніторингу якості атмосферного повітря у місті Вінниця було здійснено аналіз вхідних даних та інтерполяцію даних на прикладі показника PM2.5.

4 ЕКОНОМІЧНА ЧАСТИНА

4.1 Оцінювання комерційного потенціалу розробки

Метою проведення комерційного та технологічного аудиту є оцінювання комерційного потенціалу інформаційної технології аналізу та прогнозування якості атмосферного повітря у м. Вінниці за даними громадського моніторингу, а саме комерційного потенціалу програми, яка вибирає кращі місця для створення нових пунктів моніторингу.

Для проведення технологічного аудиту було залучено 3-х незалежних експертів Вінницького національного технічного університету кафедри системного аналізу та інформаційних технологій: к.т.н., доц. Козачко О.М., к.т.н., доц. Крижановський Є. М., к.т.н., доц. Варчук І. В. Для проведення технологічного аудиту було використано таблицю 4.1 [20] в якій за п'ятибальною шкалою використовуючи 12 критеріїв здійснено оцінку комерційного потенціалу.

Таблиця 4.1 – Рекомендовані критерії оцінювання комерційного потенціалу розробки та їх можлива бальна оцінка

Критерії оцінювання та бали (за 5-ти бальною шкалою)					
Кри-терій	0	1	2	3	4
Технічна здійсненність концепції:					
1	Достовірність концепції не підтверджена	Концепція підтверджена експертними висновками	Концепція підтверджена розрахунками	Концепція перевірена на практиці	Перевірено роботоздатність продукту в реальних умовах
Ринкові переваги (недоліки):					
2	Багато аналогів на малому ринку	Мало аналогів на малому ринку	Кілька аналогів на великому ринку	Один аналог на великому ринку	Продукт не має аналогів на великому ринку
3	Ціна продукту значно вища за ціни аналогів	Ціна продукту дещо вища за ціни аналогів	Ціна продукту приблизно дорівнює цінам аналогів	Ціна продукту дещо нижче за ціни аналогів	Ціна продукту значно нижче за ціни аналогів

Продовження таблиці 4.1

Критерії оцінювання та бали (за 5-ти бальною шкалою)					
Кри-терій	0	1	2	3	4
4	Технічні та споживчі властивості продукту значно гірші, ніж в аналогів	Технічні та споживчі властивості продукту трохи гірші, ніж в аналогів	Технічні та споживчі властивості продукту на рівні аналогів	Технічні та споживчі властивості продукту трохи кращі, ніж в аналогів	Технічні та споживчі властивості продукту значно кращі, ніж в аналогів
5	Експлуатаційні витрати значно вищі, ніж в аналогів	Експлуатаційні витрати дещо вищі, ніж в аналогів	Експлуатаційні витрати на рівні експлуатаційних витрат аналогів	Експлуатаційні витрати трохи нижчі, ніж в аналогів	Експлуатаційні витрати значно нижчі, ніж в аналогів
Ринкові перспективи					
6	Ринок малий і не має позитивної динаміки	Ринок малий, але має позитивну динаміку	Середній ринок з позитивною динамікою	Великий стабільний ринок	Великий ринок з позитивною динамікою
7	Активна конкуренція великих компаній на ринку	Активна конкуренція	Помірна конкуренція	Незначна конкуренція	Конкуренція немає
Практична здійсненність					
8	Відсутні фахівці як з технічної, так і з комерційної реалізації ідеї	Необхідно наймати фахівців або витратити значні кошти та час на навчання наявних фахівців	Необхідне незначне навчання фахівців та збільшення їх штату	Необхідне незначне навчання фахівців	Є фахівці з питань як з технічної, так і з комерційної реалізації ідеї
9	Потрібні значні фінансові ресурси, які відсутні. Джерела фінансування ідеї відсутні	Потрібні незначні фінансові ресурси. Джерела фінансування відсутні	Потрібні значні фінансові ресурси. Джерела фінансування є	Потрібні незначні фінансові ресурси. Джерела фінансування є	Не потребує додаткового фінансування
10	Необхідна розробка нових матеріалів	Потрібні матеріали, що використовуються у військово-промисловому комплексі	Потрібні дорогі матеріали	Потрібні досяжні та дешеві матеріали	Всі матеріали для реалізації ідеї відомі та давно використовуються у виробництві
11	Термін реалізації ідеї більший за 10 років	Термін реалізації ідеї більший за 5 років. Термін окупності інвестицій більше 10-ти років	Термін реалізації ідеї від 3-х до 5-ти років. Термін окупності інвестицій більше 5-ти років	Термін реалізації ідеї менше 3-х років. Термін окупності інвестицій від	Термін реалізації ідеї менше 3-х років. Термін окупності інвестицій менше 3-х років

Продовження таблиці 4.1

Критерії оцінювання та бали (за 5-ти бальною шкалою)					
Кри-терій	0	1	2	3	4
				3-х до 5-ти років	
12	Необхідна розробка регламентних документів та отримання великої кількості дозвільних документів на виробництво та реалізацію продукту	Необхідно отримання великої кількості дозвільних документів на виробництво та реалізацію продукту, що вимагає значних коштів та часу	Процедура отримання дозвільних документів для виробництва та реалізації продукту вимагає незначних коштів та часу	Необхідно тільки повідомлення відповідним органам про виробництво та реалізацію продукту	Відсутні будь-які регламентні обмеження на виробництво та реалізацію продукту

Таблиця 4.2 – Рівні комерційного потенціалу розробки

Середньоарифметична сума балів СБ, розрахована на основі висновків експертів	Рівень комерційного потенціалу розробки
0-10	Низький
11-20	Нижче середнього
21-30	Середній
31-40	Вище середнього
41-48	Високий

В таблиці 4.3 наведено результати оцінювання експертами комерційного потенціалу розробки.

Таблиця 4.3 – Результати оцінювання комерційного потенціалу розробки

Критерії	Прізвище, ініціали, посада експерта		
	Козачко О.М.	Крижановський Є.М.	Варчук І.В.
	Бали, виставлені експертами:		
1	3	2	3
2	3	2	2
3	2	4	3
4	4	4	3
5	2	3	3
6	1	2	2
7	2	1	1
8	4	3	3

Продовження таблиці 4.3

Критерії	Прізвище, ініціали, посада експерта		
	Козачко О.М.	Крижановський Є.М.	Варчук І.В.
	Бали, виставлені експертами:		
9	1	3	1
10	3	4	3
11	4	4	4
12	2	3	3
Сума балів	СБ ₁ =31	СБ ₂ =34	СБ ₃ =34
Середньоарифметична сума балів $\overline{СБ}$	$\overline{СБ} = \frac{\sum_1^3 СБ_i}{3} = \frac{31 + 34 + 34}{3} = 33$		

Середньоарифметична сума балів, розрахована на основі висновків експертів склала 33 бали, що згідно таблиці 4.2 вважається, що рівень комерційного потенціалу проведених досліджень є вище середнього.

Інформаційні технології аналізу та прогнозування якості атмосферного повітря у м. Вінниці за даними громадського моніторингу, а саме програма, яка вибирає кращі місця для створення нових пунктів моніторингу буде цікава міській громаді, а також іншим містам у яких є потреба оцінити якість повітря та визначити місця для очищення повітря та покращення навколишнього середовища в місті.

4.2 Прогнозування витрат на виконання науково-дослідної роботи

Витрати, пов'язані з проведенням науково-дослідної роботи групуються за такими статтями: витрати на оплату праці, витрати на соціальні заходи, матеріали, паливо та енергія для науково-виробничих цілей, витрати на службові відрядження, програмне забезпечення для наукових робіт, інші витрати, накладні витрати.

1. Основна заробітна плата кожного із дослідників $З_0$, якщо вони працюють в наукових установах бюджетної сфери визначається за формулою:

$$Z_0 = \frac{M}{T_p} * t \text{ (грн)}, \quad (4.1)$$

де M – місячний посадовий оклад конкретного розробника (інженера, дослідника, науковця тощо), грн.;

T_p – число робочих днів в місяці; приблизно $T_p \approx 21...23$ дні;

t – число робочих днів роботи дослідника.

Для розробки програмні засоби необхідно залучити програміста з посадовим окладом 9000 грн. Кількість робочих днів у місяці складає 40, а кількість робочих днів програміста складає 22. Зведемо сумарні розрахунки до таблиця 4.4.

Таблиця 4.4 – Заробітна плата дослідника в науковій установі бюджетної сфери

Найменування посади	Місячний посадовий оклад, грн.	Оплата за робочий день, грн.	Число днів роботи	Витрати на заробітну плату грн.
Керівник	17000	772,7	5	3864
Програмний інженер	9000	409,1	40	16364
Всього				20227

2. Розрахунок додаткової заробітної плати робітників

Додаткова заробітна плата Z_d всіх розробників та робітників, які приймали участь в розробці нового технічного рішення розраховується як 10 - 12 % від основної заробітної плати робітників.

На даному підприємстві додаткова заробітна плата начисляється в розмірі 10% від основної заробітної плати.

$$Z_d = (Z_0 + Z_p) * \frac{N_{\text{дод}}}{100\%} \quad (4.2)$$

$$З_д = 0,11 * 20227 = 2225 \text{ (грн)}.$$

3. Нарахування на заробітну плату $H_{зп}$ дослідників та робітників, які брали участь у виконанні даного етапу роботи, розраховуються за формулою:

$$H_{зп} = (З_о + З_д) * \frac{\beta}{100}, \quad (4.3)$$

де $З_о$ – основна заробітна плата розробників, грн.;

$З_д$ – додаткова заробітна плата всіх розробників та робітників, грн.;

β – ставка єдиного внеску на загальнообов'язкове державне соціальне страхування, % .

Дана діяльність відноситься до бюджетної сфери, тому ставка єдиного внеску на загальнообов'язкове державне соціальне страхування буде складати 22%, тоді:

$$H_{зп} = (20227 + 2225) * \frac{22}{100} = 4939,5 \text{ (грн)}.$$

4. Витрати на комплектуючі вироби, які використовують при виготовленні одиниці продукції, розраховуються, згідно їх номенклатури, за формулою:

$$K = \sum_{i=1}^n H_i * Ц_i * K_i, \quad (4.5)$$

де H_i – кількість комплектуючих i -го виду, шт.;

$Ц_i$ – покупна ціна комплектуючих i -го найменування, грн.;

K_i – коефіцієнт транспортних витрат (1,1...1,15).

Таблиця 4.5 – Комплектуючі, що використані на розробку

Найменування матеріалу	Ціна за одиницю, грн.	Витрачено	Вартість витраченого матеріалу, грн.
Папір	130	1	130
Ручка	15	1	15
CD-диск	20	1	20
Флешка	145	1	145
Всього			310
З врахуванням коефіцієнта транспортування			341

5. Програмне забезпечення для наукової роботи включає витрати на розробку та придбання спеціальних програмних засобів і програмного забезпечення необхідного для проведення дослідження.

Для написання магістерської роботи використовувалися інтернет середовище Kaggle та ArcGIS online, які є безкоштовними.

6. Амортизація обладнання, комп'ютерів та приміщень, які використовувались під час виконання даного етапу роботи

Дані відрахування розраховують по кожному виду обладнання, приміщенням тощо.

$$A = \frac{Ц*Т}{Т_{кор}*12}, \quad (4.6)$$

де Ц – балансова вартість даного виду обладнання (приміщень), грн.;

$T_{кор}$ – час користування;

T – термін використання обладнання (приміщень), цілі місяці.

Згідно пункта 137.3.3 Податкового кодекса амортизація нараховується на основні засоби вартістю понад 2500 грн. В нашому випадку для написання магістерської роботи використовувався персональний комп'ютер вартістю 26000 грн.

$$A = \frac{26000 \cdot 1}{2 \cdot 12} = 1083,33 \text{ (грн)}.$$

7. До статті «Паливо та енергія для науково-виробничих цілей» відносяться витрати на всі види палива й енергії, що безпосередньо використовуються з технологічною метою на проведення досліджень.

$$B_e = \sum_{i=1}^n \frac{W_{yt} \cdot t_i \cdot C_e \cdot K_{впi}}{\eta_i}, \quad (4.7)$$

де W_{yt} – встановлена потужність обладнання на певному етапі розробки, кВт;

t_i – тривалість роботи обладнання на етапі дослідження, год;

C_e – вартість 1 кВт-години електроенергії, грн;

$K_{впi}$ – коефіцієнт, що враховує використання потужності, $K_{впi} < 1$;

η_i – коефіцієнт корисної дії обладнання, $\eta_i < 1$.

Для написання магістерської роботи використовується персональний комп'ютер для якого розрахуємо витрати на електроенергію.

$$B_e = \frac{0,3 \cdot 185 \cdot 4,1 \cdot 0,5}{0,8} = 142,22 \text{ (грн)}.$$

Витрати на службові відрядження, витрати на роботи, які виконують сторонні підприємства, установи, організації та інші витрати в нашому дослідженні не враховуються оскільки їх не було.

Накладні (загальновиробничі) витрати $B_{нзв}$ охоплюють: витрати на управління організацією, оплата службових відряджень, витрати на утримання, ремонт та експлуатацію основних засобів, витрати на опалення, освітлення, водопостачання, охорону праці тощо. Накладні (загальновиробничі) витрати $B_{нзв}$ можна прийняти як (100...150)% від суми основної заробітної плати розробників та робітників, які виконували дану МКНР, тобто:

$$B_{\text{НЗВ}} = (З_0 + З_p) \cdot \frac{H_{\text{НЗВ}}}{100\%}, \quad (4.8)$$

де $H_{\text{НЗВ}}$ – норма нарахування за статтею «Інші витрати».

$$B_{\text{НЗВ}} = 20227 \cdot \frac{100}{100\%} = 20227 \text{ (грн)}.$$

Сума всіх попередніх статей витрат дає витрати, які безпосередньо стосуються даного розділу МКНР

$$B = 20227 + 2225 + 4939,5 + 341 + 1083,33 + 142,22 + 202227 = 49185,6 \text{ (грн)}.$$

Прогнозування загальних втрат ЗВ на виконання та впровадження результатів виконаної МКНР здійснюється за формулою:

$$ЗВ = \frac{B}{\eta}, \quad (4.9)$$

де η – коефіцієнт, який характеризує стадію виконання даної НДР.

Оскільки, робота знаходиться на стадії науково-дослідних робіт, то коефіцієнт $\beta = 0,9$.

Звідси:

$$ЗВ = \frac{49185,6}{0,9} = 54650,66 \text{ (грн)}.$$

4.3 Розрахунок економічної ефективності науково-технічної розробки

У даному підрозділі кількісно спрогнозуємо, яку вигоду, зиск можна отримати у майбутньому від впровадження результатів виконаної наукової

роботи. Розрахуємо збільшення чистого прибутку підприємства $\Delta\Pi_i$, для кожного із років, протягом яких очікується отримання позитивних результатів від впровадження розробки, за формулою

$$\Delta\Pi_i = \sum_1^n (\Delta\Pi_0 * N * \Pi_0 * \Delta N)_i * \lambda * \rho * \left(1 - \frac{v}{100}\right), \quad (4.10)$$

де $\Delta\Pi_0$ – покращення основного оціночного показника від впровадження результатів розробки у даному році.

N – основний кількісний показник, який визначає діяльність підприємства у даному році до впровадження результатів наукової розробки;

ΔN – покращення основного кількісного показника діяльності підприємства від впровадження результатів розробки:

Π_0 – основний оціночний показник, який визначає діяльність підприємства у даному році після впровадження результатів наукової розробки;

n – кількість років, протягом яких очікується отримання позитивних результатів від впровадження розробки:

λ – коефіцієнт, який враховує сплату податку на додану вартість. Ставка податку на додану вартість дорівнює 20%, а коефіцієнт $\lambda = 0,8333$.

ρ – коефіцієнт, який враховує рентабельність продукту. $\rho = 0,25$;

v – ставка податку на прибуток. У 2021 році – 18%.

Припустимо, що ціна за програмний продукт зросте на 500 грн. Кількість одиниць реалізованої продукції також збільшиться: протягом першого року на 45 шт., протягом другого року – на 35 шт., протягом третього року на 25 шт. Реалізація продукції до впровадження розробки складала 1 шт., а її ціна до складає 9000 грн. Розрахуємо прибуток, яке отримає підприємство протягом трьох років.

$$\begin{aligned} \Delta\Pi_1 &= [500 \cdot 1 + (9000 + 500) \cdot 45] \cdot 0,833 \cdot 0,25 \cdot \left(1 + \frac{18}{100}\right) \\ &= 73113,74 \text{ (грн)}. \end{aligned}$$

$$\begin{aligned}\Delta\Pi_2 &= [500 \cdot 1 + (9000 + 500) \cdot (45 + 35)] \cdot 0,833 \cdot 0,25 \cdot \left(1 + \frac{18}{100}\right) \\ &= 130328,14 \text{ (грн)}.\end{aligned}$$

$$\begin{aligned}\Delta\Pi_3 &= [500 \cdot 1 + (9000 + 500) \cdot (45 + 35 + 25)] \cdot 0,833 \cdot 0,25 \cdot \left(1 + \frac{18}{100}\right) \\ &= 170899,43 \text{ (грн)}.\end{aligned}$$

4.4 Розрахунок ефективності вкладених інвестицій та періоду їх окупності

Розрахуємо основні показники, які визначають доцільність фінансування наукової розробки певним інвестором, є абсолютна і відносна ефективність вкладених інвестицій та термін їх окупності.

Розрахуємо величину початкових інвестицій PV , які потенційний інвестор має вкласти для впровадження і комерціалізації науково-технічної розробки.

$$PV = k_{\text{інв}} \cdot 3B, \quad (4.11)$$

де $k_{\text{інв}}$ – коефіцієнт, що враховує витрати інвестора на впровадження науково-технічної розробки та її комерціалізацію. Це можуть бути витрати на підготовку приміщень, розробку технологій, навчання персоналу, маркетингові заходи тощо ($k_{\text{інв}} = 2 \dots 5$).

$$PV = 2 \cdot 54650,66 = 109301,33 \text{ (грн)}.$$

Розрахуємо абсолютну ефективність вкладених інвестицій $E_{\text{абс}}$ згідно наступної формули:

$$E_{abc} = (ПП - PV), \quad (4.12)$$

де ПП – приведена вартість всіх чистих прибутків, що їх отримає підприємство від реалізації результатів наукової розробки, грн.;

$$ПП = \sum_1^T \frac{\Delta\Pi_i}{(1+\tau)^t}, \quad (4.13)$$

де $\Delta\Pi_i$ – збільшення чистого прибутку у кожному із років, протягом яких виявляються результати виконаної та впровадженої НДДКР, грн.;

T – період часу, протягом якого виявляються результати впровадженої НДДКР, роки;

τ – ставка дисконтування, за яку можна взяти щорічний прогнозований рівень інфляції в країні; для України цей показник знаходиться на рівні 0,2;

t – період часу (в роках).

$$ПП = \frac{73113,74}{(1 + 0,2)^1} + \frac{130328,14}{(1 + 0,2)^2} + \frac{170899,43}{(1 + 0,2)^3} = 250793,91 \text{ (грн).}$$

$$E_{abc} = (250793,91 - 109301,33) = 141492,58 \text{ (грн).}$$

Оскільки $E_{abc} > 0$, то вкладання коштів на виконання та впровадження результатів НДДКР може бути доцільним.

Розрахуємо відносну (щорічну) ефективність вкладених в наукову розробку інвестицій E_B . Для цього користуються формулою:

$$E_B = \sqrt[T_{ж}] {1 + \frac{E_{abc}}{PV}} - 1, \quad (4.14)$$

де $T_{ж}$ – життєвий цикл наукової розробки, роки.

$$E_B = \sqrt[3]{1 + \frac{141492,58}{109301,33}} - 1 = 0,53 = 53\%.$$

Визначимо мінімальну ставку дисконтування, яка у загальному вигляді визначається за формулою:

$$\tau = d + f, \quad (4.15)$$

де d – середньозважена ставка за депозитними операціями в комерційних банках; в 2021 році в Україні $d = (0,14 \dots 0,2)$;

f – показник, що характеризує ризикованість вкладень; зазвичай, величина $f = (0,05 \dots 0,1)$.

$$\tau_{min} = 0,18 + 0,5 = 0,23.$$

Так як $E_B > \tau_{min}$ то інвестор може бути зацікавлений у фінансуванні даної наукової розробки.

Розрахуємо термін окупності вкладених у реалізацію наукового проекту інвестицій за формулою:

$$T_{ок} = \frac{1}{E_B}. \quad (4.16)$$

$$T_{ок} = \frac{1}{0,53} = 1,9 \text{ (роки)}.$$

Так як $T_{ок} \leq 3 \dots 5$ -ти років, то фінансування даної наукової розробки в принципі є доцільним.

4.5 Висновки

Було проведено оцінку комерційного потенціалу інформаційної технології аналізу та прогнозування якості атмосферного повітря у місті Вінниці, а саме програми, яка вибирає кращі місця для створення нових пунктів моніторингу за даними громадського моніторингу, який є на вище середньому рівні.

Прогнозування витрат на виконання науково-дослідної роботи по кожній з статей витрат складе 49185,6 грн. Загальна ж величина витрат на виконання та впровадження результатів даної НДР буде складати 54650,66 грн.

Вкладені інвестиції в даний проект окупляться через 1,9 роки, приведена вартість всіх чистих прибутків, що їх отримає підприємство від реалізації результатів наукової розробки склала 250793,91 грн.

ВИСНОВКИ

Проблема якості повітря вважається однією з найбільш своєчасних світових завдань. Розуміння цього вже призвело до того, що держави стали брати на себе конкретні заходи для контролю і управління якістю навколишнього середовища. Створення сіток стаціонарних станцій контролю, розробка дешевих датчиків, поліпшення супутникових технологій, встановлення постів спостереження це робить посилки для нарощування розміру та підвищення якості даних, важливих для ефективного контролю забруднення повітря.

У першому розділі було здійснено аналіз предметної області, було визначено, що таке пости спостереження та які вони бувають, визначення моніторингу атмосферного повітря та його етапи. Також, розглянуто сучасний стан моніторингу атмосферного повітря у місті Вінниця, визначено основні проблеми та проекти, що займаються встановленням постів по місту.

У другому розділі було здійснено вибір оптимальних інформаційних технологій, огляд вхідних даних, відображення постів з набором даних на карті, та аналіз показників, що визначаються постами моніторингу якості атмосферного повітря.

У третьому розділі на основі зібраних даних спостережень з постів мережі громадського моніторингу якості атмосферного повітря у місті Вінниця було здійснено аналіз вхідних даних та інтерполяцію даних на прикладі показника PM_{2.5}. Аналіз показав, що для того, аби побудувати якісну варіограму та по ній – картограму реального чи прогнозованого стану забруднення атмосферного повітря у м. Вінниці за даними громадського моніторингу, потрібно мати регулярні дані мінімум з 20 постів, а у Вінниці є поки лише 9.

Варіограми побудовані з великою похибкою, оскільки є значна дисперсія розкиду між значеннями. Однією з причин цього може бути те, що, одночасно, і дисперсія – завелика, і точок – замало. Тому розроблена методика

вибору місць для нових пунктів спостереження, яка передбачає те, що варто ставити нові пости, передусім, там, де є суттєві зміни (велика дисперсія) значень основних показників якості атмосферного повітря (мікрорайони Вишенька, Слов'янка та ін., наприклад, на вулицях Тараса Шевченка, 600-річчя та проспекті Юності), а по-друге, там – де постів замало (в першу чергу, у південній і північній частинах міста – мікрорайони Корея, П'ятничани, Старе місто та ін.).

У четвертому розділі було проведено оцінку комерційного потенціалу інформаційної технології аналізу та прогнозування якості атмосферного повітря у місті Вінниці, а саме програми, яка вибирає кращі місця для створення нових пунктів моніторингу за даними громадського моніторингу, який є на вище середньому рівні.

Прогнозування витрат на виконання науково-дослідної роботи по кожній з статей витрат складе 49185,6 грн. Загальна ж величина витрат на виконання та впровадження результатів даної НДР буде складати 54650,66 грн.

Вкладені інвестиції в даний проект окупляться через 1,9 роки, приведена вартість всіх чистих прибутків, що їх отримає підприємство від реалізації результатів наукової розробки склала 250793,91 грн.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Князюк А. О., Мокін В. Б., Мокін О. Б. ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ІНТЕРПОЛЮВАННЯ ДАНИХ ПРО ЯКІСТЬ АТМОСФЕРНОГО ПОВІТРЯ У М. ВІННИЦІ ЗА ДАНИМИ ГРОМАДСЬКОГО МОНІТОРИНГУ. *Всеукраїнська науково-практична інтернет-конференція «Молодь в науці: дослідження, проблеми, перспективи» (Вінниця, 2021-2022 рр.)* [Електронний ресурс]. URL: <https://conferences.vntu.edu.ua/index.php/mn/mn2022/paper/viewFile/14236/1204>
- 3.
2. Мокін В. Б. Автоматизована система екоінспекційного контролю стану забруднення довкілля України та викидів, скидів і відходів "Екоінспектор". Частина I. Підсистема "Викиди.": методичний посібник. Вінниця : УНІВЕРСУМ-Вінниця, 2009. 192 с.
3. Метод оцінювання параметрів стаціонарного джерела викиду на основі моделі Гауса за даними оперативного моніторингу зони розсіювання / В. Б. Мокін, Д. Ю. Дзюняк. Вінниця : Математичне моделювання в економіці, 2016. 27-35 с.
4. Мокін В. Б. Створення інформаційної системи моніторингу забруднення атмосферного повітря міста на основі технології «Інтернет речей». Мокін В. Б., Собко Б. Ю, Крижановський Є. М., Дратований М.В., Горячев Г. В.. Вінниця : Вісник Вінницького політехнічного інституту, 2017. 49-58 с.
5. Мокін В. Б., Дзюняк Д. Ю, Бондалєтов К. О., Олійник В. В.. Метод та технологія моніторингу стану атмосферного повітря за допомогою універсальної інформаційно-вимірювальної системи з використанням мобільних пристроїв. *Наукові праці Вінницького національного технічного університету*. Вінниця, 2015. № 4.[Електронний ресурс]. URL: <http://praci.vntu.edu.ua/index.php/praci/article/view/456/454>

6. Мокін В. Б. Моделювання поширення забруднювальних речовин у повітрі міста з використанням геоінформаційних технологій / В. Б. Мокін, І. В. Варчук, Вісник Вінницького політехнічного інституту. 2013. № 5. С. 13-18.

7. Мокін В. Б. Аналіз впливу капітального ремонту доріг на стан атмосферного повітря з використанням геоінформаційних технологій на прикладі м. Вінниці / В. Б. Мокін, Ю. С. Семчук, О. П. Сорочан, О. В. Риженко, Екологічна безпека та природокористування: зб. наук. праць. Київ, 2011. Вип. 7. С. 5–15.

8. Мокін В.Б. Новий метод розробки інформаційних моделей систем екологічного контролю / В. Б. Мокін, А. Р. Ящолт, Міжнародний науково-технічний журнал «Інформаційні технології та комп'ютерна інженерія». 2008. № 1(11). С. 37–43.

9. Мокін В. Б. Новий підхід щодо розробки автоматизованої геоінформаційної системи для комплексного екологічного контролю стану забруднення довкілля / В. Б. Мокін, М. П. Боцула, Є. М. Крижановський, А. Р. Ящолт, Вісник Національного технічного університету України «Київський політехнічний інститут», Серія «Хімічна інженерія, екологія та ресурсозбереження». Київ : НТУУ «КПІ». 2013. № 1 (11). С. 97-102.

10. Мокін В. Б. Системний ГІС-аналіз впливу асфальтування доріг та ремонту центрального моста на забруднення атмосферного повітря м. Вінниці у 2017 році / В. Б. Мокін, Є. М. Крижановський, А. І. Лебухорський. Збірник доповідей XLVII науково-технічної конференції професорсько-викладацького складу, співробітників та студентів університету з участю працівників науково-дослідних організацій та інженерно-технічних працівників підприємств м. Вінниці та області. Вінниця : ВНТУ, 2018. [Електронний ресурс]. URL: <https://conferences.vntu.edu.ua/index.php/all-fksa/all-fksa-2018/paper/view/4467/4420>

11. Технологія оцінювання параметрів викидів зі стаціонарних джерел за даними оперативного моніторингу / В. Мокін, Д. Дзюняк : Матеріали III Міжнародної науково-практичної конференції "Summer InfoCom Advanced

Solutions 2016", м. Київ, 1-2 грудня 2016 р., Київ : ТОВ «Інженеринг», 2016. С. 66-68.

12. Мокін В. Б. Підвищення точності моделювання забруднення атмосферного повітря міста з використанням мобільних інформаційно-вимірювальних систем / В. Б. Мокін, Г. В. Горячев, Д. Ю. Дзюняк, К. О. Бондалетов : XLV Регіональна науково-технічна конференція професорсько-викладацького складу, співробітників та студентів університету з участю працівників науково-дослідних організацій та інженерно-технічних працівників підприємств м. Вінниці та області, електронне наукове видання матеріалів конференції, м. Вінниця, 23-24 березня 2016. [Електронний ресурс]. URL: <http://conferences.vntu.edu.ua/index.php/all-ebmd/all-ebmd-2016/paper/view/465/866>

13. Мокін В. Б. Практична реалізація мобільної аналітичної комп'ютерної системи моніторингу стану атмосферного повітря з підсистемою веб-аналізу та виведення даних на геопортал / В. Б. Мокін, Д. Ю. Дзюняк, К. О. Бондалетов, Г. В. Горячев, В. В. Боцян : V-й Всеукраїнський з'їзд екологів з міжнародною участю (Екологія/Ecology-2015), 23-26 вересня, 2015. Збірник наукових праць. Вінниця: ТОВ «Нілан-ЛТД», 2015. С. 106.

14. Мокін В. Б. Технологія оперативного екологічного моніторингу стану повітряного середовища та пересувних джерел його забруднення / В. Б. Мокін, Д. Ю. Дзюняк, К. О. Бондалетов : XIV Міжнародна науково-практична конференція: Сучасні інформаційні технології управління екологічною безпекою, природокористуванням, заходами в надзвичайних ситуаціях : 5-9 жовтня 2015 р. Київ, 2015. С. 126-130.

15. Боцян В. В. Моніторинг стану забруднення атмосферного повітря міста оксидом вуглецю з використанням мобільної аналітичної інформаційно-вимірювальної системи / В. В. Боцян, В. Б. Мокін : Молодь в технічних науках: дослідження, проблеми, перспективи : матеріали Міжнародної науково-практичної Інтернет-конференції, 23-26 квітня 2015 р. Вінниця, 2015. С. 31-33.

16. Мокін В. Б. Розробка АРМ геоінформаційної системи екологічного моніторингу міста Кривий Ріг / В. Б. Мокін, О. В. Гавенко : "Наука. Молодь. Екологія" (в рамках I Всеукраїнського молодіжного з'їзду екологів з міжнародною участю) : матеріали Міжнародної науково-практичної конференції, 21-23 травня 2014 р. Житомир, 2014. С. 29-33.

17. Мокін В. Б. Новий підхід до розробки інформаційних моделей автоматизованих систем екоінспекційного контролю / В. Б. Мокін, А. Р. Ящолт : Зб. матеріалів X Міжнародної науково-технічної конференції «Системний аналіз та інформаційні технології». 20-24 травня 2008 р. Київ, 2008. 385 с.

18. Мокін В. Б. Розробка імітаційної моделі та комп'ютерної програми для автоматизації процесу еколого-економічного планування управління природокористуванням у регіоні із дотриманням вимог екологічної безпеки / В. Б. Мокін, А. Р. Ящолт, О. В. Крилик, Б. Ю. Собко, С. О. Довгополук : XIV Міжнародна науково-практична конференція: Сучасні інформаційні технології управління екологічною безпекою, природокористуванням, заходами в надзвичайних ситуаціях, 5-9 жовтня 2015 р. Київ, 2015. С. 135-139.

19. Мокін В. Б., Горячев Г. В., Дзюняк Д. Ю. Інформаційна ГІС-інтегрована технологія оцінювання параметрів викидів речовин за даними оперативного моніторингу забруднення атмосферного повітря : монографія. Вінниця: ВНТУ, 2018. 100 с.

20. Козловський В. О., Лесько О. Й., Кавецький В. В. Методичні вказівки до виконання економічної частини магістерських кваліфікаційних робіт : уклад. Вінниця : ВНТУ, 2021. 42 с.

Додаток А

Міністерство освіти і науки України
Вінницький національний технічний університет
Факультет комп'ютерних систем і автоматики

ЗАТВЕРДЖУЮ

Завідувач кафедри САІТ

_____ д.т.н., проф. Мокін В. Б.

«__» _____ 2021 р.

ТЕХНІЧНЕ ЗАВДАННЯ

на магістерську кваліфікаційну роботу

«ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ АНАЛІЗУ ТА ПРОГНОЗУВАННЯ
ЯКОСТІ АТМОСФЕРНОГО ПОВІТРЯ У МІСТІ ВІННИЦІ ЗА
ДАНИМИ ГРОМАДСЬКОГО МОНІТОРИНГУ (КОМПЛЕКСНА
МКР). ЧАСТИНА ІІ. ІНТЕРПОЛЮВАННЯ ДАНИХ ТА ВИБІР
МІСЦЬ ДЛЯ НОВИХ ПУНКТИВ»

08-53.МКР.002.02.000.ТЗ

Керівник: д.т.н., проф. каф. САІТ

_____ Мокін О. Б.

«__» _____ 2021 р.

Розробила: студентка гр. ЗІСТ-20м

_____ Князюк А. О .

«__» _____ 2021 р.

Вінниця 2021

1. Підстава для проведення робіт

Підставою для виконання роботи є наказ № __ по ВНТУ від «__» _____ 2021 р., та індивідуальне завдання на МКР, затверджене протоколом № __ засідання кафедри САІТ від «__» _____ 2021 р.

2. Джерела розробки:

- Мокін В. Б. Моделювання поширення забруднювальних речовин у повітрі міста з використанням геоінформаційних технологій / В. Б. Мокін, І. В. Варчук, Вісник Вінницького політехнічного інституту. 2013. № 5. С. 13-18.
- Мокін В. Б., Горячев Г. В., Дзюняк Д. Ю. Інформаційна ГІС-інтегрована технологія оцінювання параметрів викидів речовин за даними оперативного моніторингу забруднення атмосферного повітря : монографія. Вінниця: ВНТУ, 2018. 100 с.

3. Мета і призначення роботи:

Поліпшення точності побудови картограм реального чи прогнозованого стану забруднення атмосферного повітря у м. Вінниці за даними громадського моніторингу шляхом розроблення методики вибору місць для нових пунктів спостереження.

4. Вихідні дані для проведення робіт:

Набір даних моніторингу атмосферного повітря по 9 постах наданими ресурсами Eco City та SaveEcoBot.

5. Методи дослідження:

- розвідувальний аналіз;
- інтерполювання даних.

6. Етапи роботи і терміни їх виконання:

1. Аналіз предметної області _____ – _____
2. Розробка методу дослідження даних _____ – _____
3. Розробка моделей виявлення закономірностей _____ – _____
4. Оформлення пояснювальної записки. _____ – _____

7. Очікувані результати та порядок реалізації:

Визначити найкраще місце розташування для встановлення нових постів моніторингу атмосферного повітря у місті Вінниця.

8. Вимоги до розробленої документації

Пояснювальна записка оформлена у відповідності до вимог «Методичних вказівок до виконання та оформлення магістерських кваліфікаційних робіт для студентів спеціальності 126 – «Інформаційні системи та технології» денної форми навчання».

9. Порядок приймання роботи

Публічний захист «__» _____ .2021 р.

Початок розробки «__» _____ 2021 р.

Граничні терміни виконання МКР «__» _____ 2021 р.

Розробила студентка групи ЗІСТ-20м _____ Князюк А.О.

Додаток Б

Протокол перевірки кваліфікаційної роботи

Назва роботи: «Інформаційна технологія аналізу та прогнозування якості атмосферного повітря у місті Вінниці за даними громадського моніторингу (комплексна МКР). Частина II. Інтерполювання даних та вибір місць для нових пунктів»

Тип роботи: магістерська кваліфікаційна робота

Підрозділ: кафедра САІТ

Науковий керівник: Мокін О.Б. д.т.н., проф. каф. САІТ

Показники звіту подібності

Unicheck	
Оригінальність	90,3 %
Схожість	9,7 %

Аналіз звіту подібності (відмітити потрібне)

- Запозичення, виявлені у роботі, оформлені коректно і не містять ознак плагіату.
- Виявлені у роботі запозичення не мають ознак плагіату, але їх надмірна кількість викликає сумніви щодо цінності роботи і самостійності її автора. Роботу направити на доопрацювання.
- Виявлені у роботі запозичення є недобросовісними і мають ознаки плагіату та/або в ній містяться навмисні спотворення тексту, що вказують на спроби приховування недобросовісних запозичень.

Заявляю, що ознайомена з повним звітом подібності, який був згенерований системою щодо роботи

Автор _____ Князюк А.О.

Опис прийнятого рішення

Робота допускається до захисту

Особа, відповідальна за перевірку _____ Жуков С. О.

Додаток В
Лістинг програми

```

indicator_name = 'PM2.5'
time_interval='H'
type_agg='mean'
datetime_analysis = '2021-11-27 09:00:00'
stations_about = pd.read_csv('../input/air-quality-
monitoring/saveecobot_city_about_stations.csv', header=0, sep=';')
stations_about = stations_about[stations_about['locality']=='Vinnytsia
city'].reset_index(drop=True)
stations_about = stations_about.fillna(-1)
stations_about
stations_about['id_station'] = "SaveEcoBot_" +
stations_about['id_saveecobot'].astype('str')
stations_about['id_ecocity'] = stations_about['id_ecocity'].astype('int')
stations_about.loc[stations_about['network']=='Eco-City', 'id_station'] =
"EcoCity_" + stations_about['id_ecocity'].astype('int').astype('str')
stations_about
stations_about = stations_about[['id_station', 'lat', 'lng']]
stations_about
ecocity_stations_about = pd.read_csv('../input/air-quality-monitoring-from-
ecocity/ecocity_about_stations_2021.csv', header=0, sep=';')
ecocity_stations_about
ecocity_stations_about_city =
ecocity_stations_about[ecocity_stations_about['locality']=='Vinnytsia
city'].reset_index(drop=True)
ecocity_stations_about_city['id_ecocity'] =
ecocity_stations_about_city['id_ecocity'].astype('int')
ecocity_stations_about_city

```

```

ecocity_stations_about_city['id_station'] = "EcoCity_" +
ecocity_stations_about_city['id_ecocity'].astype('str')
ecocity_stations_about_city.loc[ecocity_stations_about_city['network']=='Eco-City',
'id_station'] = "EcoCity_" +
ecocity_stations_about_city['id_ecocity'].astype('str')
ecocity_stations_about_city
ecocity_stations_about_city = ecocity_stations_about_city[['id_station', 'lat',
'lng']]
ecocity_stations_about_city
ecocity_stations_about_city_list =
ecocity_stations_about_city['id_station'].unique().tolist()
ecocity_stations_about_city_list
stations_about =
stations_about[~stations_about['id_station'].isin(ecocity_stations_about_city
_list)]
stations_about
stations_about = pd.concat([ecocity_stations_about_city, stations_about],
ignore_index=True)
stations_about
num = len(stations_about)
stations_about.loc[num, 'id_station'] = 'VinnCHM_1'
stations_about.loc[num, 'lat'] = 49.241962
stations_about.loc[num, 'lng'] = 28.478482
stations_about.loc[num+1, 'id_station'] = 'VinnCHM_2'
stations_about.loc[num+1, 'lat'] = 49.231392
stations_about.loc[num+1, 'lng'] = 28.518974
stations_about
stations_about_list = stations_about['id_station'].tolist()
stations_about_list

```

```

st_par = pd.read_csv('../input/air-quality-
monitoring/stations_parameters_prediction.csv', header=0, sep=';')
st_par
def f(st_par, id_station, indicator_name, t):

    df = st_par[st_par['id_station']==id_station].reset_index(drop=True)
    k1 = float(df.loc[0, indicator_name+"_k1"])
    k0 = float(df.loc[0, indicator_name+"_k0"])
    a1 = float(df.loc[0, indicator_name+"_a1"])
    b1 = float(df.loc[0, indicator_name+"_b1"])

    return k0 + k1*t + a1*np.sin(t) + b1*np.cos(t)
    res = pd.DataFrame(columns=['id_station', 't', 'y'])
res.loc[0,'t']=1
res = pd.DataFrame(columns=['id_station', 'indicator_name', 't', 'y'])
w = 0
for id_station in stations_about_list:
    for indicator_name in ['PM2.5', 'PM10']:
        for t in range(10):
            res.loc[w, 'id_station'] = id_station
            res.loc[w, 'indicator_name'] = indicator_name
            res.loc[w, 't'] = t
            res.loc[w, 'y'] = f(st_par, id_station, indicator_name, t) +
np.random.normal(0, .1, 1)[0]
        w += 1
res
stations_about
m = folium.Map(location=[49.2355, 28.474], zoom_start=13, tiles="Stamen
Terrain")

```

```

for i in range(len(stations_about)):
    marker = folium.Marker(
        location=[float(stations_about.loc[i, 'lat']), float(stations_about.loc[i,
'lng'])],
        popup=folium.Popup(max_width=450)
    )
    marker.add_to(m)

```

m

```

url = (
    "https://raw.githubusercontent.com/python-
visualization/folium/master/examples/data"
)
data = json.loads(requests.get(f"{url}/vis1.json").text)

```

```

m = folium.Map(location=[49.2355, 28.474], zoom_start=13)

```

```

for i in range(len(stations_about)):
    marker = folium.Marker(
        location=[float(stations_about.loc[i, 'lat']), float(stations_about.loc[i,
'lng'])],
        popup=folium.Popup(max_width=450).add_child(
            folium.Vega(data, width=450, height=450) # Still need to replace
data to res for given id_station and indicator
        ),
    )
    marker.add_to(m)

```

m

Додаток Г

ІЛЮСТРАТИВНА ЧАСТИНА

ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ АНАЛІЗУ ТА ПРОГНОЗУВАННЯ
ЯКОСТІ АТМОСФЕРНОГО ПОВІТРЯ У МІСТІ ВІННИЦІ ЗА ДАНИМИ
ГРОМАДСЬКОГО МОНІТОРИНГУ (КОМПЛЕКСНА МКР). ЧАСТИНА ІІ.
ІНТЕРПОЛЮВАННЯ ДАНИХ ТА ВИБІР МІСЦЬ ДЛЯ НОВИХ ПУНКТИВ

Виконала: студентка гр. 2ІСТ-20м

_____ Князюк А. О.

« ___ » _____ 2021 р.

Керівник: д.т.н., проф. каф. САІТ

_____ Мокін О. Б.

« ___ » _____ 2021 р.

Нормоконтроль: к.т.н., доцент

_____ Жуков С. О.

« ___ » _____ 2021 р.

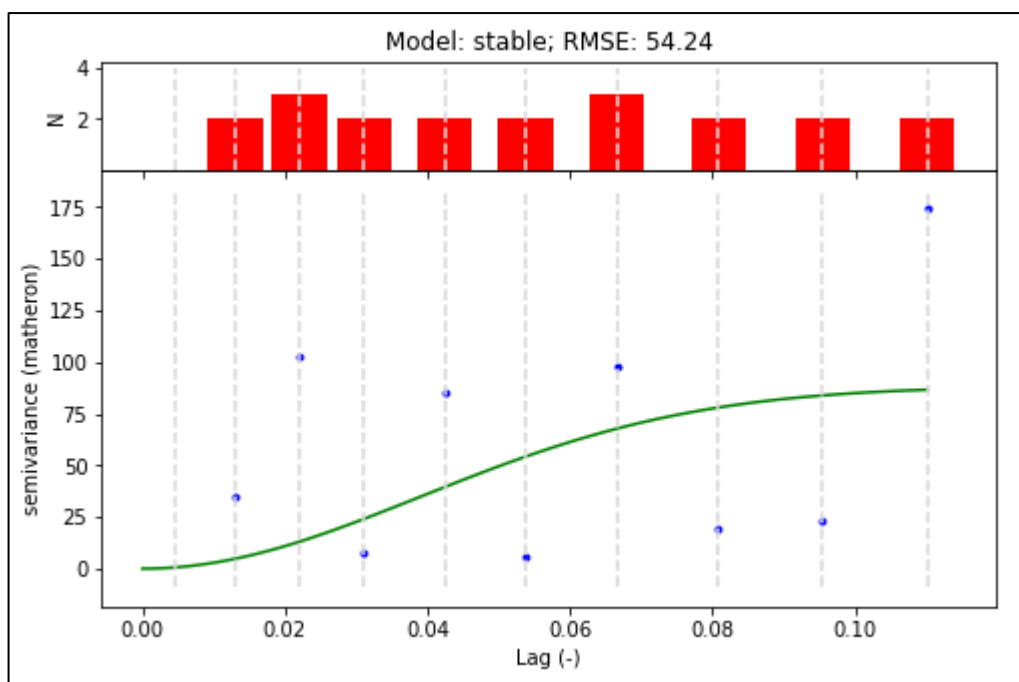


Рисунок Г.1 – Сферична варіограма по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 26.11.2021 о 00:00:00

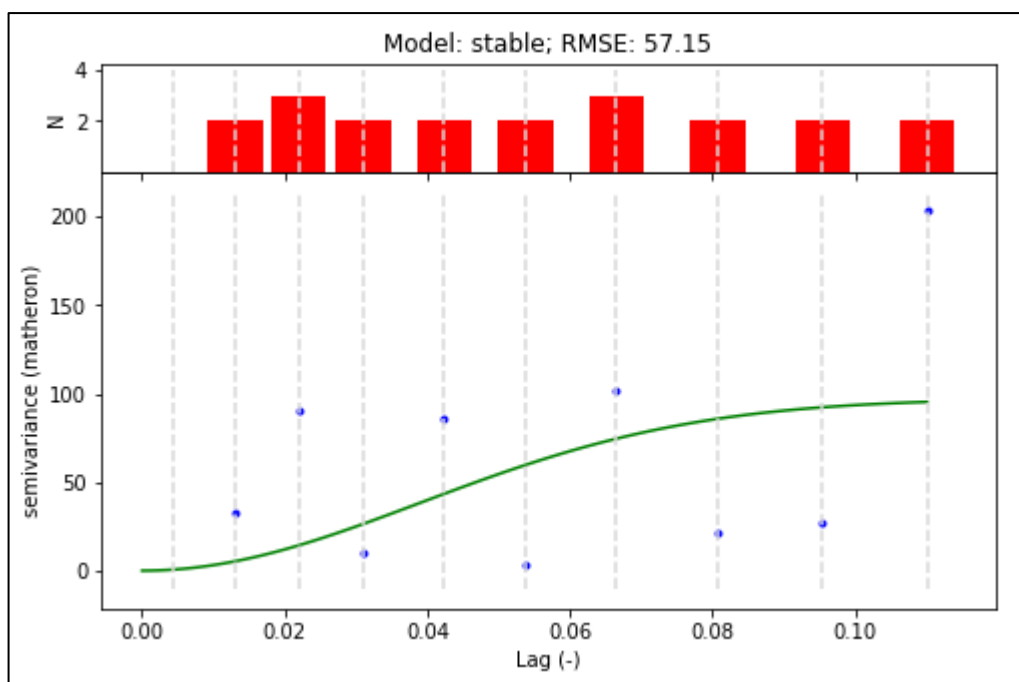


Рисунок Г.2 – Сферична варіограма по показнику пилу PM_{2.5} (концентрація частинок розміром 2,5 мкм) за 26.11.2021 о 01:00:00

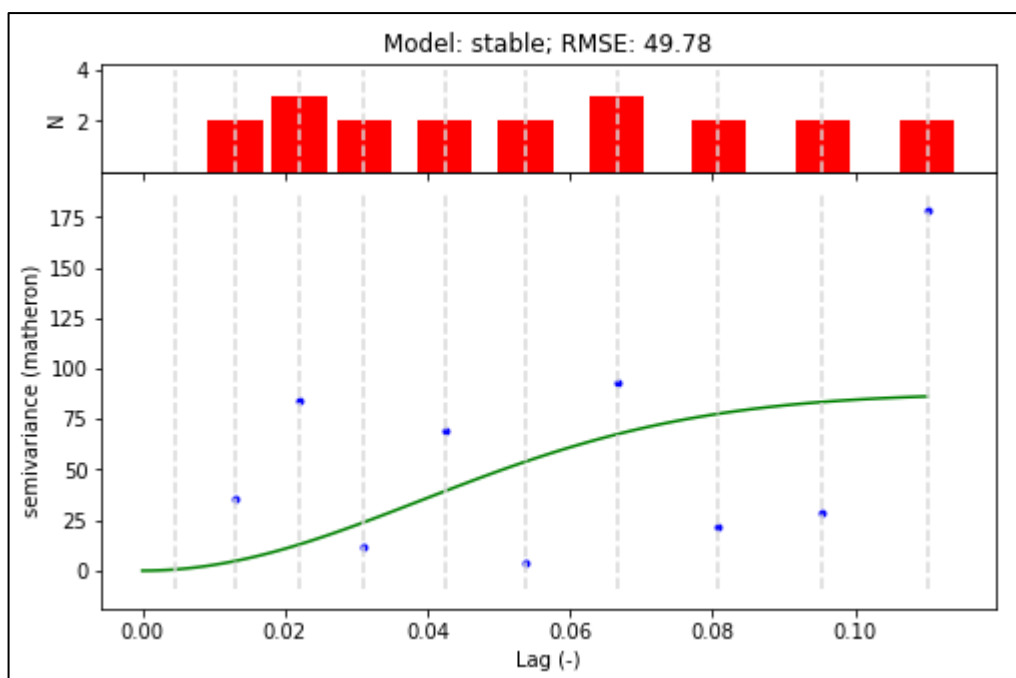


Рисунок Г.3 – Сферична варіограма по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 26.11.2021 о 02:00:00

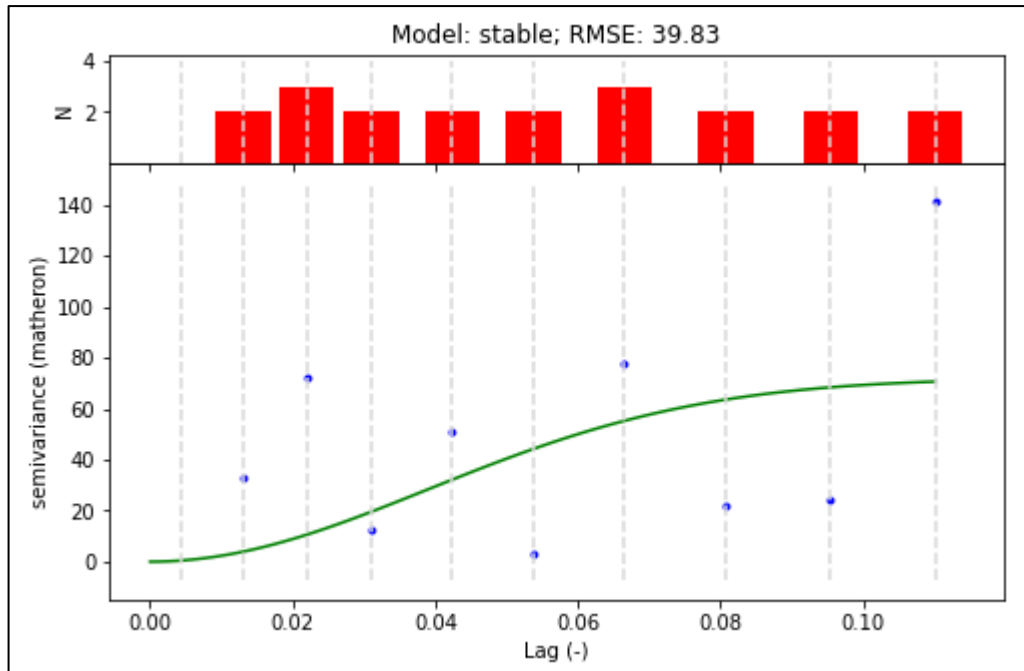


Рисунок Г.4 – Сферична варіограма по показнику пилу PM_{2.5} (концентрація частинок розміром 2,5 мкм) за 26.11.2021 о 03:00:00

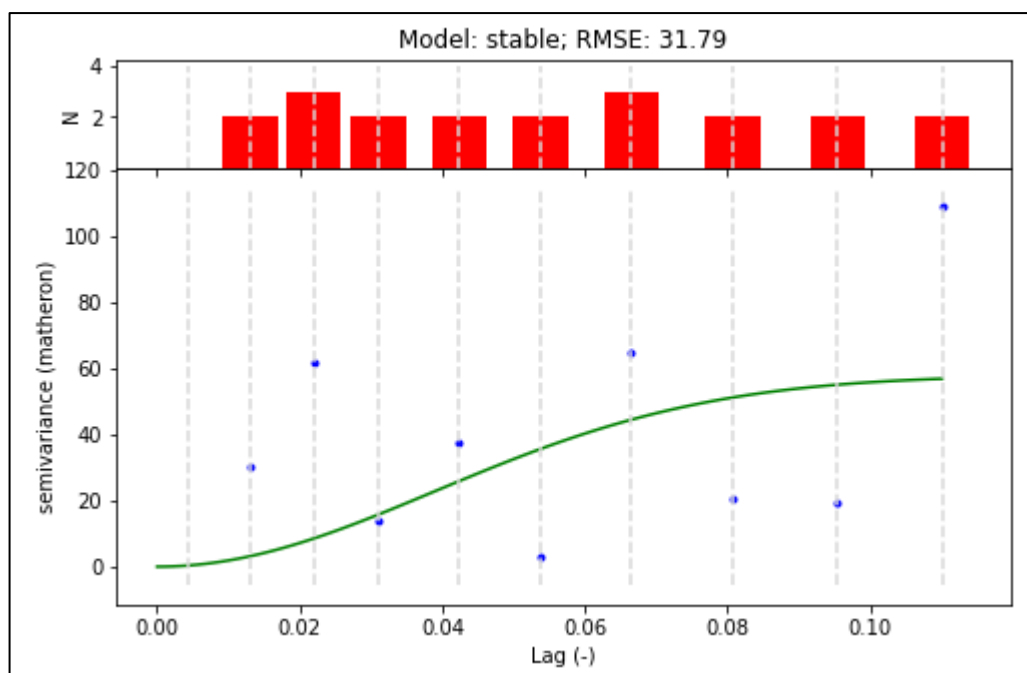


Рисунок Г.5 – Сферична варіограма по показнику пилу PM_{2.5} (концентрація частинок розміром 2,5 мкм) за 26.11.2021 о 04:00:00

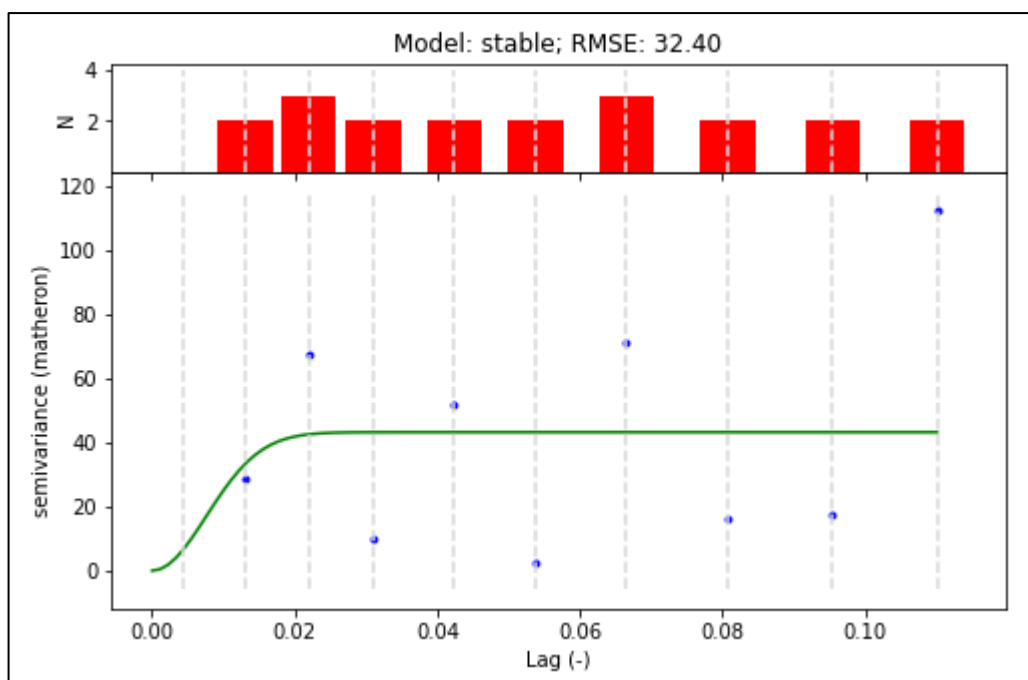


Рисунок Г.6 – Сферична варіограма по показнику пилу PM_{2.5} (концентрація частинок розміром 2,5 мкм) за 26.11.2021 о 05:00:00

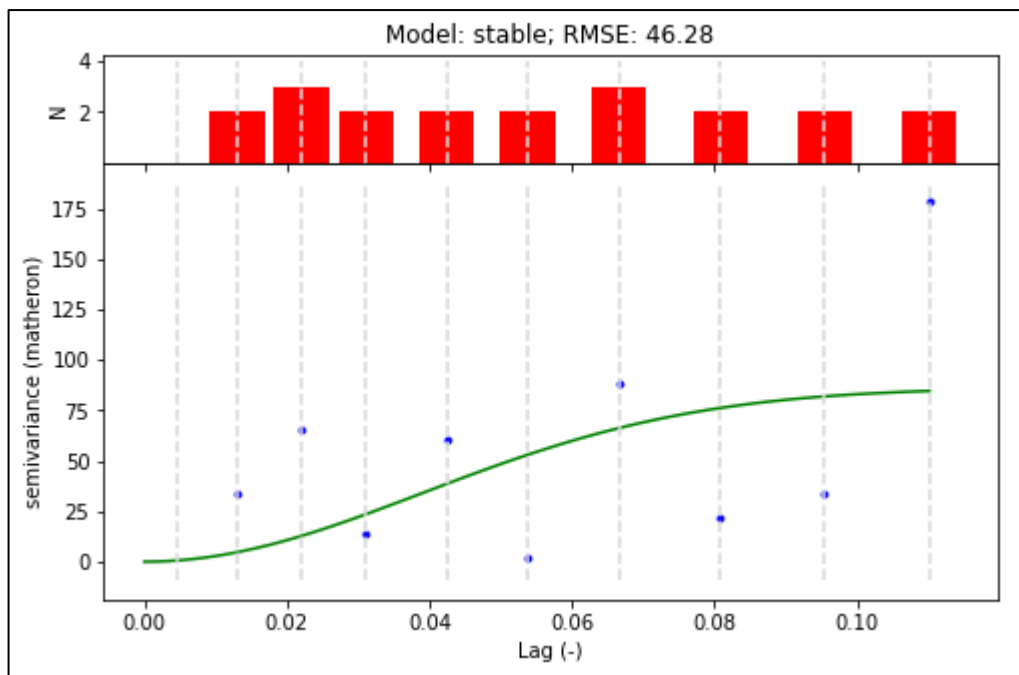


Рисунок Г.7 – Сферична варіограма по показнику пилу PM2.5 (концентрація частинок розміром 2,5 мкм) за 26.11.2021 о 06:00:00