

Вінницький національний технічний університет
Факультет інтелектуальних інформаційних технологій та автоматизації
Кафедра системного аналізу та інформаційних технологій

МАГІСТЕРСЬКА КВАЛІФІКАЦІЙНА РОБОТА

на тему:

«Інформаційна технологія геостатистичного аналізу радіаційного забруднення в Європі»

Виконав: студент 2 курсу, групи 2ІСТ-23м
спеціальності 126 – «Інформаційні
системи та технології»

Богдан Богдан ОПАНАСЮК

Керівник: д.т.н., проф. каф. САІТ

Олександр Олександр МОКІН
«05» 12 2024 р.

Опонент: д.т.н., проф. каф. АІТ

Ярослав Ярослав ІВАНЧУК
«08» 12 2024 р.

Допущено до захисту

Завідувач кафедри САІТ

Віталій д.т.н., проф. Віталій МОКІН
«06» 12 2024 р.

Вінниця ВНТУ – 2024 рік

Вінницький національний технічний університет
Факультет інтелектуальних інформаційних технологій та автоматизації
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. Віталій МОКІН

«17» 09 2024 року

ЗАВДАННЯ
НА МАГІСТЕРСЬКУ КВАЛІФІКАЦІЙНУ РОБОТУ СТУДЕНТУ
Опанасюку Богдану Михайловичу

1. Тема роботи: «Інформаційна технологія геостатистичного аналізу радіаційного забруднення в Європі»,
керівник роботи: Олександр МОКІН д.т.н., професор. каф. САІТ,
затверджені наказом ВНТУ від «17» 09 2024 року № 310
2. Термін подання студентом роботи «01» 12 2024 року
3. Вихідні дані до роботи:
Датасет Kaggle «Chernobyl Chemical Radiation / CSV / Country Data»
4. Зміст текстової частини:
 - Аналіз сучасного стану моніторингу атмосферного повітря
 - Вибір сучасних підходів до обробки і аналізу даних
 - Розроблення інформаційної технології геостатистичного аналізу радіаційного забруднення в Європі;
 - Економічна частина.
5. Перелік ілюстративного матеріалу:
 - Спрощений приклад кригінг-інтерполяції
 - Метод інтерполяції Кригінг
 - Картографічне відображення вхідних даних
 - UML Component Diagram системи;
 - Діаграма розсіювання даних після передбачення допоміжних даних
 - Порівняння основних моделей побудови варіограм
 - Результат геостатистичного аналізу основними популярними моделями

6. Консультанти розділів МКР

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
4	Олександр ЛЕСЬКО, к.е.н., проф. каф. ЕПВМ	 09.11.24	 30.11.24

7. Дата видачі завдання «07» 09 2024 року

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва та зміст етапу	Термін виконання		Примітка
		початок	закінчення	
1	Аналіз предметної області	30.09.	10.10	визн
2	Вибір оптимальних інформаційних технологій	11.10	31.10	визн
3	Розроблення інформаційної технології	01.11	15.11	визн
4	Економічна частина	16.11	25.11	визн
5	Оформлення матеріалів до захисту МКР	26.11	30.11	визн

Студент

Керівник роботи

Богдан ОПАНАСЮК

Олександр МОКІН

АНОТАЦІЯ

УДК 004.09

Опанасюк Б.М. Інформаційна технологія геостатистичного аналізу радіаційного забруднення в Європі. Магістерська кваліфікаційна робота зі спеціальності 126 – інформаційні системи та технології, освітньо-професійна програма – інформаційні технології аналізу даних та зображень. Вінниця: ВНТУ, 2024. 109 с.

На укр. мові. Бібліогр.: 32 назв; рис.: 35; табл.: 12.

У ході дослідження проаналізовано можливість побудови інтерполяційних картограм стану радіаційного забруднення у країнах Європи за даними державного моніторингу. Об'єктом досліджень є процес аналізу рівня радіаційного забруднення країн Європи після катастрофи на ЧАЕС. Удосконалено інформаційну технологію геостатистичного аналізу радіаційного забруднення в Європі, що дозволило вдосконалити карту забруднення на території Європи після аварії на ЧАЕС.

У розділі економічної частини розглянуто питання про доцільність розробки та впровадження інформаційної технології геостатистичного аналізу радіаційного забруднення в Європі.

Ілюстративна частина складається з 7 плакатів.

Ключові слова: моніторинг радіаційного забруднення, інтерполяція, геостатистичний аналіз, моніторинг, аналіз даних.

ABSTRACT

Opanasiuk B.M. Information technology of geostatistical analysis of radiation pollution in Europe. Master's qualification work in specialty 126 - information systems and technologies, educational and professional program - information technologies of data and image analysis. Vinnytsia: VNTU, 2024. 109 p.

In Ukrainian. Bibliography: 32 titles; Fig.: 35; Table: 12.

The study analyzed the possibility of constructing interpolation cartograms of the state of radiation pollution in European countries according to state monitoring data. The object of research is the process of analyzing the level of radiation pollution in European countries after the Chernobyl disaster. The information technology of geostatistical analysis of radiation pollution in Europe has been improved, which allowed to improve the map of pollution in Europe after the Chernobyl accident.

The economic section considers the feasibility of developing and implementing information technology for geostatistical analysis of radiation contamination in Europe.

The illustrative part consists of 7 posters.

Keywords: radiation pollution monitoring, interpolation, geostatistical analysis, monitoring, data analysis.

ЗМІСТ

ВСТУП	8
1 АНАЛІЗ СУЧАСНОГО СТАНУ МОНІТОРИНГУ АТМОСФЕРНОГО ПОВІТРЯ.....	11
1.1 Аналіз предметної області.....	11
1.2 Аспекти моніторингу радіаційного фону у країнах Європи.	12
1.3 Історія та еволюція інформаційних технологій у моніторингу радіаційного забруднення.....	15
1.3.1 Вдосконалення методів інтерполяції та аналізу даних.....	16
1.3.2 Використання геоінформаційних систем	20
1.3.3 Застосування хмарних технологій та Big Data у радіаційному моніторингу	24
1.4 Висновки	28
2 ВИБІР СУЧАСНИХ ПІДХОДІВ ДО ОБРОБКИ ТА АНАЛІЗУ ДАНИХ..	29
2.1 Вибір рішень для розв’язання поставленого завдання.....	29
2.1.1 ArcGIS Online	31
2.1.2 Python як основний інструмент аналізу даних	33
2.1.3 Kaggle як платформа для виконання обчислень.....	34
2.2 Дослідження інтерполяції даних як явища.....	35
2.2.1 Інтерполяція даних за допомогою методу Kriging.....	36
2.2.2 Інструменти інтерполяції.....	37
2.2.3 Порівняння методів інтерполяції.....	39
2.3 Висновки	44
3 РОЗРОБЛЕННЯ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ.....	45
3.1 Підготовка вхідних даних	45
3.2 Кластеризація та інтерполяція даних	47
3.3 Геостатистичний аналіз за допомогою SciKit-GStat.....	51
3.4 Висновки	55
4 ЕКОНОМІЧНА ЧАСТИНА.....	58
4.1 Оцінювання наукового ефекту.....	58
4.2 Розрахунок витрат на здійснення науково-дослідної роботи.....	62
4.2.1 Витрати на оплату праці.....	62

4.2.2 Відрахування на соціальні заходи.....	65
4.2.2 Сировина та матеріали	65
4.2.3 Розрахунок витрат на комплектуючі.....	67
4.2.5 Спецустаткування для наукових (експериментальних) робіт	67
4.2.6 Програмне забезпечення для наукових робіт.....	68
4.2.7 Амортизація обладнання, програмних засобів та приміщень	69
4.2.8 Паливо та енергія для науково-виробничих цілей	71
4.2.9 Службові відрядження	72
4.2.10 Витрати на роботи, які виконують сторонні підприємства, установи і організації.....	72
4.2.11 Інші витрати.....	73
4.2.12 Накладні (загальновиробничі) витрати	73
4.3 Оцінювання важливості та наукової значимості науково-дослідної роботи.....	74
4.4 Висновок.....	76
ВИСНОВКИ	77
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	79
Додаток А (обов'язковий) Технічне завдання	84
Додаток Б (обов'язковий) Протокол перевірки кваліфікаційної роботи на наявність текстових запозичень	87
Додаток В (довідковий) Лістинг програми.....	88
Додаток Г (обов'язковий) Ілюстративна частина.....	90

ВСТУП

Актуальність теми. Радіаційне забруднення становить серйозну загрозу для здоров'я людей і навколишнього середовища. Після Чорнобильської аварії в 1986 році та інших радіаційних інцидентів, питання моніторингу радіаційного фону є вкрай важливим для забезпечення екологічної безпеки населення Європи.

Геостатистичний аналіз дозволяє більш точно та ефективно моделювати просторовий розподіл радіаційного забруднення. Інформаційні технології, що використовуються в геостатистиці, дають змогу інтегрувати різні джерела даних, автоматизувати обробку великих обсягів інформації та підвищити точність прогнозів.

Наявність точних і актуальних даних про радіаційне забруднення дозволяє органам влади та міжнародним організаціям приймати обґрунтовані рішення щодо заходів безпеки, планування евакуацій, розробки нормативних актів і проведення екологічних оцінок.

Інформаційні технології, що використовуються для геостатистичного аналізу, мають системи управління базами даних, ГІС (геоінформаційні системи), машинне навчання та інші сучасні технології обробки даних. Їхнє використання сприяє розвитку нових методів аналізу та покращенню існуючих, що важливо для багатьох сфер науки та техніки.

Проблема радіаційного забруднення не обмежується національними кордонами, тому міжнародна співпраця у цій сфері є вкрай важливою. Використання уніфікованих інформаційних технологій для аналізу даних сприяє обміну інформацією між країнами та міжнародними організаціями, що дозволяє ефективніше координувати зусилля для розв'язання проблем радіаційного забруднення.

Мета і завдання роботи. Метою магістерської кваліфікаційної роботи є підвищення ефективності аналізу радіаційного забруднення в Європі шляхом створення інформаційної технології на основі геостатистичних методів.

Розробка технології визначення нових місць передбачає виконання таких завдань як:

- Аналіз сучасних інформаційних технологій, що використовуються для оброблення та аналізу просторових даних, вибір середовища та оптимальних технологій;
- Вибір джерел даних про радіаційне забруднення в Європі (дані про фон радіації, вимірювання концентрацій радіонуклідів тощо), лінійна інтерполяція даних;
- Оброблення та підготовка даних для подальшого геостатистичного аналізу;
- Застосування методів інтерполяції та інших геостатистичних методів для створення карт забруднення;
- Інтерпретація отриманих результатів та їх візуалізація у вигляді карт, графіків та інших наочних матеріалів.

Об'єктом дослідження магістерської кваліфікаційної роботи є процес моніторингу та аналізу якості радіаційного забруднення країн Європи після аварії на ЧАЕС.

Предметом дослідження магістерської кваліфікаційної роботи є методи та інструменти геостатистичного аналізу, а також їх застосування для оцінювання, моделювання та прогнозування просторового розподілу радіаційного забруднення на території Європи.

Методи дослідження. У дослідженнях використовувались сучасні інформаційні технології, методи геостатистичного аналізу, геоінформаційні технології та методи пакету програм ArcGIS, методи бібліотек на Python для аналізу даних та інші інструменти, які дозволяють ефективно розв'язувати поставлені задачі та досягати мету дослідження.

Наукова новизна. Дістала подальший розвиток інформаційна технологія

аналізу радіаційного забруднення в Європі, за рахунок використання геостатистичного методу кригінгу для аналізу просторових даних на основі моделей «matern» та «stable», що дозволяє покращити точність та ефективність моніторингу, прогнозування та управління радіаційними ризиками в Європі

Апробація результатів магістерської кваліфікаційної роботи.

Результати роботи подані на всеукраїнську науково-практичну інтернет-конференцію «Молодь в науці: дослідження, проблеми, перспективи» (Вінниця, 2024-2025 рр.).

Публікації результатів магістерської кваліфікаційної роботи.

Опубліковано тези на Всеукраїнській науково-практичній інтернет-конференції «Молодь в науці: дослідження, проблеми, перспективи» (Вінниця, 2024-2025 рр.) [21].

1 АНАЛІЗ СУЧАСНОГО СТАНУ МОНІТОРИНГУ АТМОСФЕРНОГО ПОВІТРЯ

1.1 Аналіз предметної області

Аварія на Чорнобильській атомній електростанції у 1986 році призвела до значного радіаційного забруднення великих територій, що стало серйозною екологічною та медичною проблемою сучасних країн Європи. З того часу було створено різноманітні системи моніторингу для контролю рівня радіаційного забруднення та захисту населення.

Сучасний стан моніторингу базується на використанні новітніх технологій і методів, зокрема геоінформаційних систем (ГІС), які активно застосовуються для збору, зберігання, аналізу та візуалізації просторових даних про радіаційне забруднення. Інтерактивні картографічні інструменти дозволяють в реальному часі відстежувати рівні радіації на різних територіях, а супутники, обладнані датчиками, забезпечують можливість вимірювати рівень радіації та оцінювати стан забруднення великих територій. Дані з автоматизованих систем моніторингу, встановлених на стаціонарних та мобільних платформах, передаються в реальному часі до центральних баз даних для подальшого аналізу [3].

Інформаційні технології, такі як Big Data та аналітика, дозволяють обробляти великі обсяги інформації з різних джерел, забезпечуючи більш точний аналіз радіаційного забруднення. Хмарні платформи сприяють зберіганню та обробці даних, що дозволяє швидко доступати до інформації та ділитися нею між різними установами та країнами. Міжнародна співпраця відіграє важливу роль у моніторингу радіаційного забруднення, зокрема через організації, такі як МАГАТЕ, які координують моніторинг та обмін даними, а також програми, як от ЕМЕР (European Monitoring and Evaluation Programme), що забезпечують спільний моніторинг та оцінку забруднення [4]. Спільні міжнародні проекти

спрямовані на розвиток нових технологій та методів для моніторингу та зниження наслідків радіаційного забруднення, а співпраця між країнами дозволяє обмінюватися досвідом та передовими практиками.

Однак, існують певні виклики, з якими стикаються системи моніторингу. Технічні виклики включають необхідність постійного оновлення та модернізації обладнання, а також складнощі у зборі та обробці великих обсягів даних з різних джерел. Організаційні виклики полягають у координації між різними організаціями та країнами для забезпечення ефективного моніторингу, а також у забезпеченні фінансування для підтримки та розвитку цих систем. Екологічні та соціальні виклики включають вплив радіаційного забруднення на навколишнє середовище та здоров'я населення, а також необхідність інформування та захисту населення від наслідків радіаційного забруднення.

Сучасний стан моніторингу радіаційного забруднення після аварії на ЧАЕС характеризується широким використанням новітніх технологій та методів, активною міжнародною співпрацею, але також стикається з низкою технічних, організаційних та екологічних викликів. Інформаційні технології та геостатистичний аналіз відіграють ключову роль у покращенні ефективності та точності моніторингу, що сприяє зменшенню впливу радіаційного забруднення на людей і довкілля.

1.2 Аспекти моніторингу радіаційного фону у країнах Європи.

Основні проблеми з моніторингом радіаційного фону країн Європи охоплюють кілька технічних, організаційних та екологічних аспектів:

а) Технічні проблеми:

1) Застаріле обладнання: У багатьох країнах використовується застаріле обладнання для моніторингу, що може не забезпечувати необхідну точність і чутливість для виявлення низьких рівнів радіації.

2) Недостатня щільність мережі моніторингу: В деяких

регіонах Європи мережа моніторингових станцій є недостатньо щільною, що призводить до прогалин у даних і неможливості вчасно виявляти локальні забруднення.

3) Інтероперабельність: Різні країни використовують різні стандарти та протоколи збору даних, що ускладнює інтеграцію та порівняння інформації на міжнародному рівні.

b) Організаційні проблеми:

1) Координація між країнами: Відсутність ефективної координації між країнами Європи може призводити до недостатнього обміну даними та інформацією, що ускладнює комплексний аналіз радіаційного фону.

2) Відсутність єдиної стратегії: Відсутність загальноєвропейської стратегії з моніторингу та управління радіаційним забрудненням може призводити до нерівномірного рівня захисту та підготовки до можливих інцидентів.

3) Проблеми з обробкою та аналізом даних: Недостатня кількість спеціалістів і обмежені можливості для обробки великих обсягів даних можуть ускладнювати своєчасний аналіз та інтерпретацію результатів моніторингу.

c) Екологічні та соціальні проблеми:

1) Різні рівні забруднення: Високі рівні забруднення в окремих регіонах, особливо поблизу промислових зон і колишніх ядерних об'єктів, вимагають особливого підходу та додаткових ресурсів для моніторингу.

2) Вплив зміни клімату: Різка клімату може впливати на розповсюдження радіоактивних речовин, що вимагає постійного оновлення методологій моніторингу та адаптації систем до нових умов.

Проте багато організацій займається фінансуванням та контролем станцій виміру радіаційного фону у Європі [7].

Зокрема JOINT RESEARCH CENTRE [7], фінансований Європейським Союзом веде проект Radioactivity Environmental Monitoring, який був заснований одразу після вибуху на ЧАЕС у 1986р. На рисунку 1.1 можна побачити дані за період 12-19 червня 2024 року надані даним дослідницьким центром.

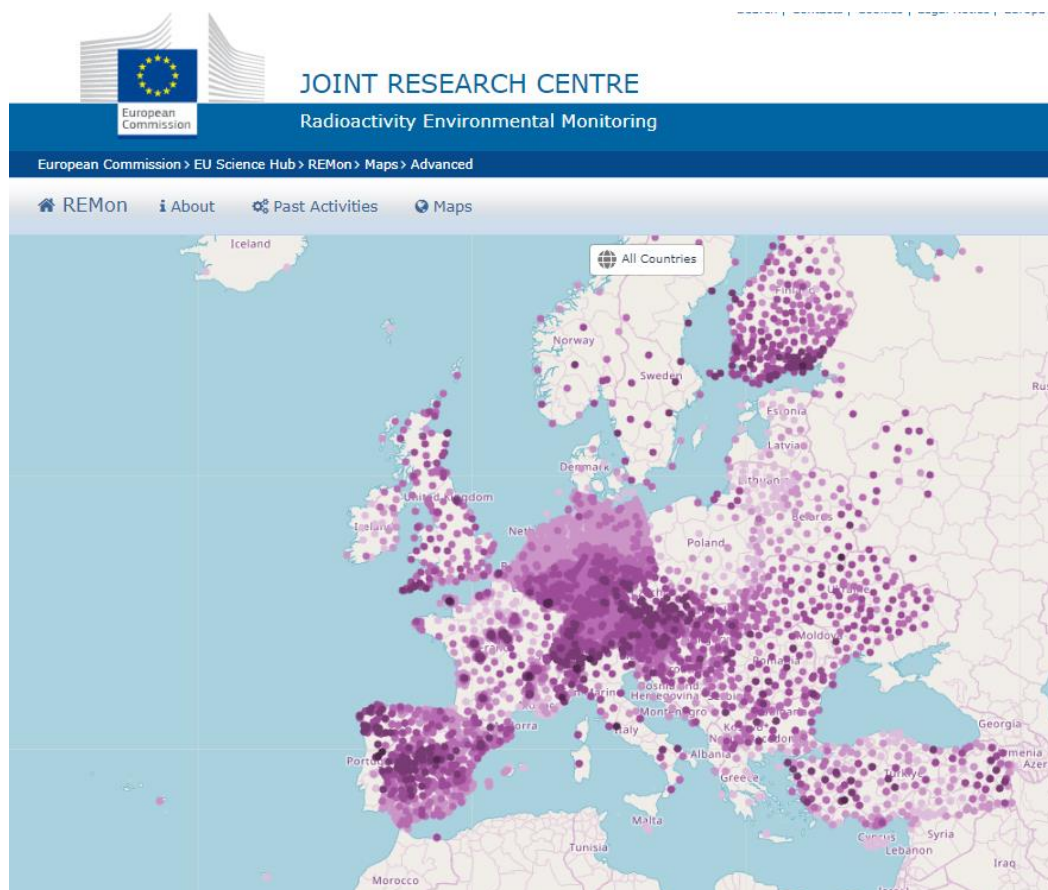


Рисунок 1.1 – Сайт Об'єданого дослідного центру ЄС

На сайті можливо не лише переглянути дані про радіаційний фон протягом певного періоду, а й ознайомитись з доступними деталями усіх екологічних проектів які проводив Об'єднаний дослідний центр протягом усього свого існування.

Також можна відзначити проект SaveEcoBot що збирає радіологічні дані на території України. На рисунку 1.2 зображена мапа з показниками радіаційного фону усіх публічно доступних станцій виміру.

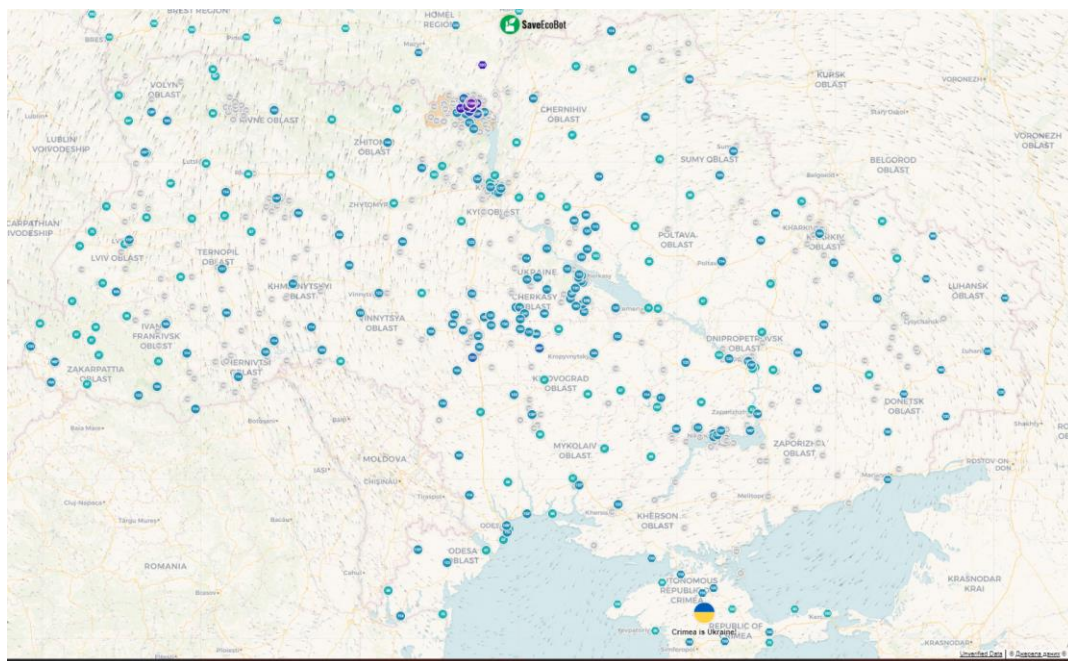


Рисунок 1.2 – Радіологічна мапа SaveEcoBot

Загалом моніторинг та прогнозування радіаційного фону є прогресивною та перспективною сферою досліджень, оскільки щороку з'являється дедалі більше підприємств та об'єктів ядерної енергетики, тому аби не повторити приклад катастроф на ЧАЕС та Фукусіма, людство повинне зосередитись на забезпеченні всіх необхідних безпекових норм щодо експлуатації таких засобів. Аналітичні розробки в цій галузі можуть використовувати як різні енергетичні та екологічні організації, так і державні інституції.

1.3 Історія та еволюція інформаційних технологій у моніторингу радіаційного забруднення

Моніторинг радіаційного забруднення став критично важливим завданням після великих ядерних катастроф, таких як аварія на Чорнобильській атомній електростанції (1986 рік) та аварія на Фукусімі (2011 рік). У перші роки після Чорнобиля застосовувалися переважно ручні методи збору даних про радіаційне забруднення, що вимагало значних людських і матеріальних ресурсів. Поступово, з розвитком інформаційних технологій (ІТ), цей процес автоматизувався і набув нових можливостей, що дозволило більш точно і швидко

аналізувати дані та прогнозувати наслідки забруднень.

Ранні системи моніторингу радіаційного фону ґрунтувалися на ручних методах вимірювання рівнів радіації, що вимагало використання портативних дозиметрів та підготовки фахівців для проведення вимірювань. Ці методи були дуже точними, але вони не могли охопити велику територію або оперативно забезпечити швидкий доступ до даних у реальному часі.

З появою перших комп'ютерних систем, як наприклад **GAMMA-3** (рис. 1.3), були автоматизовані процеси збору і обробки інформації. Ці системи збирали дані з радіаційних детекторів, розташованих на стаціонарних пунктах спостереження, що суттєво підвищило ефективність моніторингу.



Рисунок 1.3 - Детектор радіоактивного випромінювання GAMMA-3

1.3.1 Вдосконалення методів інтерполяції та аналізу даних

Одним з найважливіших досягнень у сфері моніторингу радіаційного забруднення стало впровадження методів інтерполяції, таких як інтерполяція методом зворотного зважування відстаней (IDW) та кригінг [5, 6]. Ці методи дозволяли більш точно інтерполювати рівні радіації в місцях, де вимірювання не проводилися, що дозволило зменшити кількість необхідних пунктів моніторингу та отримувати більш точні результати, і тому не безпричинно їх можна вважати найбільш важливими методами в сучасних системах моніторингу радіаційного забруднення, особливо в умовах обмеженої кількості точок спостереження. З появою складніших алгоритмів обробки даних, зокрема в галузі геостатистики, суттєво покращилися можливості інтерполяції та аналізу радіаційного фону,

навіть у випадках, коли вимірювання проводяться нерегулярно або з великими пропусками.

Метод зворотного зважування відстаней (IDW) є одним з найпростіших та найпоширеніших способів інтерполяції. Основний принцип методу полягає в тому, що найближчі до досліджуваної точки значення матимуть більший вплив на результуюче значення, ніж більш віддалені точки. Це означає, що IDW підходить для ситуацій, де існує пряма залежність між величиною радіаційного фону та відстанню до джерела забруднення.

Таблиця 1.1 - Переваги та недоліки IDW інтерполяції

Переваги	Недоліки
Простота у реалізації та розрахунках.	Неточність у випадках різко змінних даних або наявності складного рельєфу, оскільки метод не враховує варіацій в просторовій кореляції даних.
Відносно швидке виконання, навіть при великих обсягах даних.	Чутливість до кількості точок вимірювання: чим менше доступних точок, тим нижча точність інтерполяції.
Можливість використання на територіях з обмеженою кількістю точок спостереження.	

На рисунку 1.4 можна побачити приклад того, як метод IDW застосовується для інтерполяції даних радіаційного фону на основі кількох стаціонарних точок моніторингу.

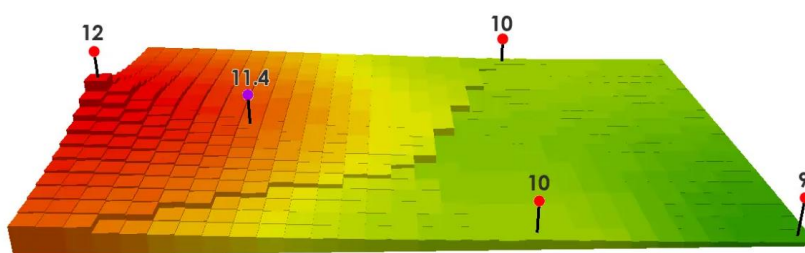


Рисунок 1.4.- Приклад інтерполяції методом зворотного зважування відстаней

Кригінг-метод є більш складним та точним порівняно з IDW, оскільки він враховує не лише відстань до точок вимірювання, а й просторову кореляцію між цими точками. Він базується на статистичній моделі, що дозволяє оцінити не лише значення у невідомій точці, але й невизначеність цього передбачення. Кригінг широко використовується в геостатистиці для інтерполяції значень у нерегулярних точках на основі відомих даних з найближчих точок.

Таблиця 1.2. Переваги та недоліки кригінг-інтерполяції

Переваги	Недоліки
Висока точність інтерполяції завдяки врахуванню просторової кореляції	Складність у реалізації: метод вимагає налаштування моделей та підготовки великих обсягів даних для точного аналізу.
Можливість оцінки невизначеності результату, що дозволяє визначити ступінь надійності прогнозу	Велика потреба у обчислювальних ресурсах, що може стати проблемою для обробки великих територій у режимі real-time.
Гнучкість: кригінг може застосовуватися на складних географічних територіях, зокрема на гірських районах або біля водойм	

Метод кригінг був вперше розроблений для потреб гірничої справи, але з часом його адаптували для багатьох галузей, включаючи екологію та моніторинг радіаційного забруднення. Однією з ключових особливостей кригінгу є використання варіограм, що дозволяють моделювати просторову структуру даних.

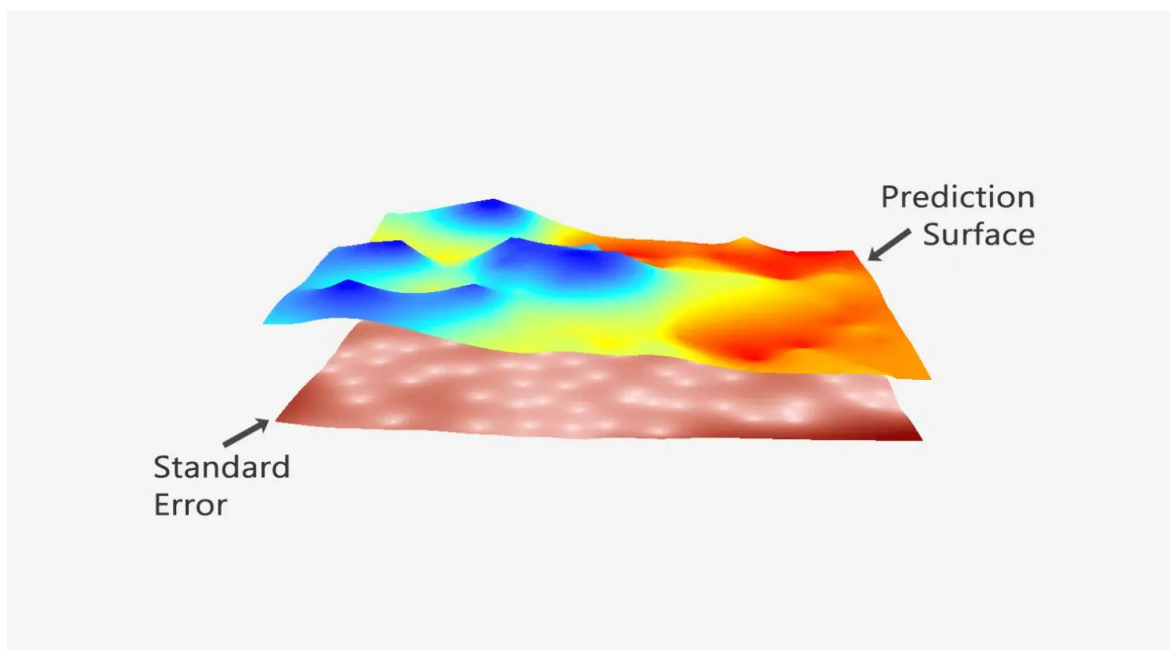


Рисунок 1.5.- Приклад застосування методу кригінг

Кригінг та IDW є двома найпоширенішими методами інтерполяції в галузі радіаційного моніторингу. Вибір методу залежить від кількості доступних точок даних, просторових особливостей території та вимог до точності результатів. У випадках, коли потрібно отримати грубіші результати з меншою кількістю обчислень, IDW може бути ефективнішим. Проте для складних територій з великою кількістю даних та необхідністю точної оцінки, кригінг є більш надійним методом. Порівняльна таблиця обох методів зображена нижче.

Таблиця 1.3. Порівняння методів IDW та кригінг

Параметр	IDW	Кригінг
Точність	Середня	Висока
Обчислювальна складність	Низька	Висока
Прогноз невизначеності	Відсутній	Є
Використання просторової кореляції	Ні	Так
Придатність для великих територій	Обмежена	Висока

Вдосконалення методів інтерполяції та аналізу даних суттєво підвищило точність прогнозів радіаційного забруднення. Сучасні інформаційні технології, зокрема використання ГІС та алгоритмів машинного навчання, дозволили інтегрувати великі обсяги даних та покращити моделювання радіаційних загроз [15]. Це є критично важливим для забезпечення ефективного контролю над екологічною ситуацією і зменшення ризиків для здоров'я людей.

1.3.2 Використання геоінформаційних систем

Геоінформаційні системи (ГІС) є одним з найважливіших досягнень в області екологічного моніторингу, і вони відіграють ключову роль у зборі, зберіганні, обробці та візуалізації просторових даних. Впровадження ГІС у системи моніторингу радіаційного забруднення дозволило значно підвищити ефективність процесів аналізу та прогнозування. У сучасному контексті ГІС забезпечують не лише інтеграцію даних з різних джерел, але й доступ до інформації в реальному часі, що особливо важливо для прийняття рішень в умовах радіаційної небезпеки.

Перші застосування геоінформаційних систем для моніторингу радіаційного забруднення почалися після катастрофи на Чорнобильській АЕС у 1986 році. У той час науковці та організації потребували інструментів, які могли б обробляти великі обсяги даних і представляти їх у вигляді картографічних моделей для оцінки рівня радіації на великих територіях.

Однією з перших ГІС-платформ, які використовувалися для цього, була

GRASS GIS – вільна система для аналізу просторових даних, що розроблялася під егідою армії США. Завдяки її можливостям з обробки даних та візуалізації результатів, система була адаптована для використання в цивільних цілях, зокрема в моніторингу екологічних та радіаційних ризиків [6].

Поступово ГІС розвивалися, і з появою інтернет-технологій почали з'являтися платформи, які дозволяли виводити дані в реальному часі. Сьогодні такі системи, як ArcGIS, стали золотим стандартом для моніторингу радіаційного фону та інших екологічних параметрів.

Сучасні ГІС-системи забезпечують значно більше, ніж просто візуалізацію даних на карті. Вони дозволяють інтегрувати різні типи даних – від показників радіаційного фону, зібраних за допомогою наземних станцій, до супутникових даних, що відображають стан атмосфери або ґрунту на великих територіях [9]. До основних функцій ГІС можна віднести:

1. Збір даних: ГІС дозволяє автоматично збирати дані з різних джерел – стаціонарних датчиків, мобільних моніторингових платформ, супутників тощо. Це забезпечує можливість отримання комплексної інформації для аналізу стану забруднення.
2. Аналіз даних: Використовуючи ГІС, можна застосовувати методи інтерполяції (кригінг, IDW), що дозволяє отримувати точні прогнози рівня забруднення навіть у тих місцях, де немає точних вимірювань.
3. Візуалізація даних: Карти та 3D-моделі, які генеруються за допомогою ГІС, дозволяють наочно представляти дані про радіаційний фон, а також аналізувати їх з просторового погляду.
4. Прогнозування: На основі даних, зібраних і оброблених у ГІС, можна моделювати можливі шляхи поширення радіації та прогнозувати її рівень у майбутньому. Це дозволяє своєчасно вживати запобіжних заходів та планувати евакуацію в разі радіаційної небезпеки.

Одним із найяскравіших прикладів використання ГІС для моніторингу радіаційного забруднення є платформа ArcGIS. Ця платформа широко використовується як міжнародними організаціями, такими як МАГАТЕ, так і

національними екологічними агенціями для моніторингу рівня радіації. На рисунку 1.6 зображений приклад візуалізації даних з використанням засобів ArcGIS радіаційного забруднення після катастрофи на АЕС Фукусіма.

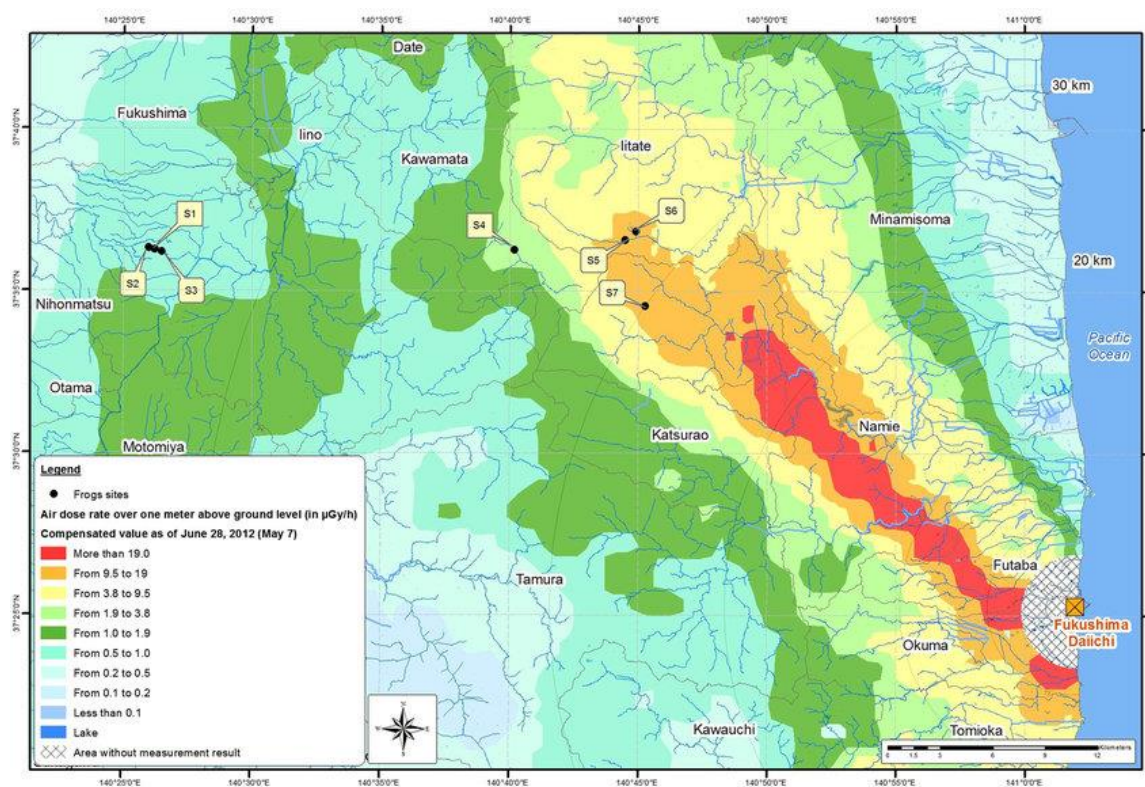


Рисунок 1.6. - Візуалізація даних радіаційного забруднення за допомогою ArcGIS [9]

За допомогою ArcGIS можливо створювати інтерактивні карти, на яких відображається рівень радіації в реальному часі. Це особливо важливо у випадках раптових радіаційних викидів, коли необхідно швидко отримати інформацію про зони забруднення та надати її органам влади для прийняття відповідних заходів.

ГІС активно використовуються в міжнародних проєктах, присвячених екологічному моніторингу. Наприклад, JOINT RESEARCH CENTRE (Об'єднаний дослідницький центр ЄС) веде проєкт Radioactivity Environmental Monitoring (REMON), який був створений після Чорнобильської катастрофи. Цей проєкт використовує ГІС для збору і аналізу даних про радіаційний фон у країнах Європи.

Одним із ключових напрямів розвитку ГІС-систем є інтеграція з хмарними

платформами. Використання хмарних технологій, таких як Google Cloud або AWS, дозволяє значно підвищити ефективність зберігання та обробки даних у ГІС. Це особливо важливо для моніторингу радіаційного фону на великих територіях, де обсяги даних можуть бути надзвичайно великими.

Хмарні ГІС-системи забезпечують можливість доступу до даних у будь-який момент часу з будь-якої точки світу, що дозволяє організувати глобальний моніторинг радіаційного фону та забезпечити обмін інформацією між країнами та організаціями. Більш того, вони дозволяють зберігати історичні дані, що важливо для аналізу довгострокових тенденцій змін радіаційного фону.

Розвиток геоінформаційних систем продовжує відкривати нові можливості для покращення моніторингу радіаційного забруднення (рис. 1.7). Серед основних напрямків розвитку можна виділити:

1. Автоматизація процесів збору та обробки даних: Використання штучного інтелекту та машинного навчання в поєднанні з ГІС дозволить автоматизувати процеси аналізу даних та прогнозування радіаційного забруднення.
2. Інтеграція з мобільними платформами: ГІС-системи можуть бути інтегровані з мобільними пристроями, що дозволить громадянам самостійно вимірювати рівень радіації та передавати ці дані в загальну систему моніторингу.
3. Розширення можливостей візуалізації: Завдяки розвитку 3D-моделювання та технологій доповненої реальності, ГІС дозволятимуть створювати детальні моделі забруднених територій, що полегшить прийняття рішень на місцевому рівні.

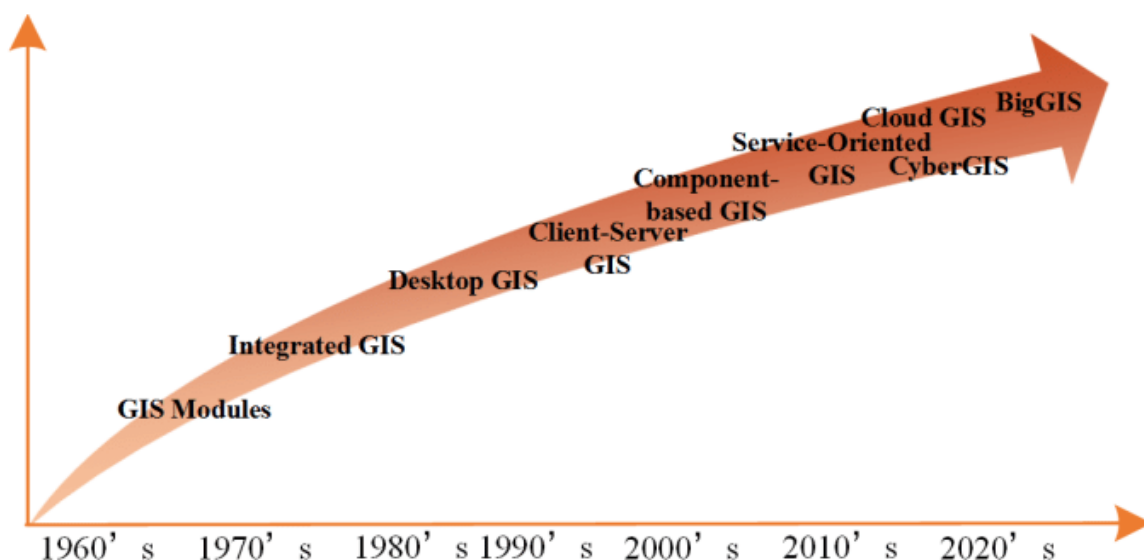


Рисунок 1.7. Прогноз розвитку ГІС [9]

Геоінформаційні системи стали невід'ємною частиною сучасного моніторингу радіаційного забруднення. Їх застосування дозволяє значно підвищити точність та ефективність аналізу даних, забезпечуючи комплексний підхід до моніторингу та прогнозування радіаційних загроз. Розвиток ГІС у майбутньому, зокрема за рахунок інтеграції з хмарними технологіями та штучним інтелектом, відкриє нові можливості для захисту населення і навколишнього середовища від радіаційного забруднення.

1.3.3 Застосування хмарних технологій та Big Data у радіаційному моніторингу

З розвитком сучасних інформаційних технологій і зростанням обсягів даних, що генеруються різними системами моніторингу, виникла необхідність у нових підходах до їх обробки та зберігання. Використання хмарних технологій та методів обробки Big Data значно змінило спосіб організації моніторингу радіаційного забруднення, забезпечивши швидкий доступ до інформації, можливість масштабування, а також більш ефективне управління ресурсами [10].

Насамперед з часом дедалі яскравіше почали з'являтися виклики традиційних систем зберігання та обробки даних, адже до появи хмарних платформ та Big Data аналітики, тогочасні системи моніторингу покладалися виключно на локальні сервери та програмне забезпечення для збору і обробки

даних про радіаційний фон. Цей підхід мав певні обмеження:

- Обмежені можливості для масштабування: після великих радіаційних катастроф, як, наприклад, аварії на Чорнобильській АЕС, кількість даних, що потрібно було обробити, значно перевищувала попередньо розраховані можливості традиційних серверів. Вимірювання проводилися на численних моніторингових станціях, і управління цими даними вимагало чималих ресурсів.
- Повільна обробка даних: Зростання обсягів даних про радіаційний фон уповільнювало процес обробки і аналізу, що почало прямим чином впливати на оперативність прийняття рішень.
- Проблеми з безпекою даних: через активне використання локальних дата-центрів зберігання даних існував ризик втрати інформації внаслідок збоїв серверів або кібератак. Також доступ до даних часто був обмежений певним колом користувачів, що ускладнювало співпрацю між різними організаціями і країнами.

Хмарні технології забезпечують можливість зберігання і обробки величезних обсягів даних, отриманих з різних джерел, включаючи наземні станції, супутники та мобільні платформи. Ось основні переваги хмарних платформ для моніторингу радіаційного забруднення:

- Масштабованість та гнучкість. Однією з ключових переваг хмарних технологій є можливість масштабування. На відміну від традиційних серверів, хмарні платформи дозволяють збільшувати або зменшувати обсяг доступних обчислювальних ресурсів залежно від кількості даних, що обробляються. Це означає, що в разі радіаційних інцидентів, коли кількість даних різко збільшується, хмарні платформи можуть миттєво адаптуватися до нових потреб, забезпечуючи безперервний процес моніторингу.
- Доступ до даних у реальному часі. Хмарні платформи дозволяють збирати та аналізувати дані у реальному часі. Це надзвичайно важливо у випадках, коли необхідно оперативно реагувати на зміни рівня радіації, наприклад, під час аварій або інших надзвичайних ситуацій. Вся

інформація від різних моніторингових станцій, включаючи мобільні та стаціонарні джерела, передається до хмарних сховищ у реальному часі, що дає можливість швидко оцінити поточну ситуацію і приймати обґрунтовані рішення.

– Зберігання великих обсягів даних. Завдяки хмарним технологіям стало можливим зберігати величезні обсяги даних про радіаційний фон, що генеруються різними джерелами. Наприклад, дані супутникового моніторингу, що охоплюють великі території, можуть легко зберігатися у хмарних сховищах без необхідності купівлі та обслуговування дорогих серверів.

– Міжнародна співпраця та доступність даних. Хмарні платформи полегшують співпрацю між різними країнами та організаціями. Дані про радіаційний фон можуть бути доступні у реальному часі для науковців, урядових установ та міжнародних організацій, що сприяє обміну інформацією і забезпечує більш ефективну координацію заходів із захисту населення. Це особливо важливо для таких організацій, як МАГАТЕ або Joint Research Centre (JRC), які працюють над координацією моніторингу радіаційних загроз у Європі та світі.

– Безпека даних. Однією з головних переваг хмарних платформ є висока безпека зберігання даних. Використання хмарних сховищ дозволяє захистити інформацію від втрат у випадку технічних збоїв або природних катастроф. Крім того, більшість провайдерів хмарних сервісів забезпечують високий рівень шифрування даних та інші заходи безпеки, що знижує ризик несанкціонованого доступу до конфіденційної інформації або втрати даних.

Неможливо також обійти стороною використання засобів Big Data – технології обробки неосяжних масивів даних, що забезпечує можливість аналізу великих і різномірних джерел інформації. [10] У радіаційному моніторингу, де дані надходять з численних джерел (супутники, стаціонарні та мобільні станції, датчики тощо), обробка Big Data слугує критично важливим компонентом для

ефективного керування інформацією. До основних його переваг для використання у подібних дослідженнях можемо віднести:

1. Аналіз та прогнозування: Завдяки Big Data можливо аналізувати величезні обсяги даних і виявляти закономірності, які раніше могли залишатися непоміченими. Це дозволяє прогнозувати можливі радіаційні викиди, виявляти джерела забруднення та оцінювати вплив на довкілля та населення.

2. Інтеграція різних типів даних: Big Data дозволяє інтегрувати різні типи даних — від вимірювань радіаційного фону до метеорологічних даних і соціальних мереж. Це забезпечує більш комплексний підхід до моніторингу та прогнозування радіаційних загроз.

3. Машинне навчання: Технології машинного навчання, що використовуються в Big Data, дозволяють автоматизувати аналіз даних і знаходити складні залежності між змінними. Це дає можливість більш точно передбачати поведінку радіоактивних речовин в атмосфері або на поверхні Землі.

4. Швидкість обробки: У порівнянні з традиційними підходами до аналізу даних, Big Data забезпечує значно швидшу обробку інформації, що дозволяє оперативно реагувати на зміни рівня радіації.

5. Візуалізація даних: На основі оброблених даних Big Data можливо створювати наочні та деталізовані карти радіаційного забруднення, що дозволяють виявляти критичні зони та проводити аналіз у динаміці.

Підбиваючи підсумки, можемо сказати що хмарні технології та Big Data вважаються істинно невід’ємно важливими компонентами сучасних систем моніторингу радіаційного забруднення. Вони забезпечують можливість обробляти великі обсяги даних у реальному часі, підвищують точність прогнозування та сприяють міжнародній співпраці в боротьбі з радіаційними загрозами. Завдяки швидкості обробки, можливостям масштабування та інтеграції різних джерел даних, ці технології значно підвищили ефективність радіаційного моніторингу та захисту населення. У майбутньому використання хмарних платформ та Big Data тільки зростатиме, надаючи ще більше

можливостей для збереження безпеки навколишнього середовища.

1.4 Висновки

Основні проблеми з моніторингом радіаційного фону в країнах Європи пов'язані з технічними обмеженнями, організаційними труднощами та екологічними й соціальними викликами. Вирішення цих проблем потребує інтегрованого підходу, що включає модернізацію обладнання, підвищення рівня координації між країнами, збільшення фінансування, а також підвищення обізнаності населення та адаптацію до змінних умов навколишнього середовища.

Сучасні інформаційні технології, такі як хмарні платформи, геоінформаційні системи (ГІС), Big Data та методи машинного навчання, значно покращують процеси моніторингу радіаційного забруднення. Завдяки їхньому впровадженню системи моніторингу стали більш точними, оперативними та здатними до обробки величезних обсягів даних у реальному часі. У ході цього дослідження було проаналізовано вплив цих технологій на збирання, обробку і візуалізацію даних про радіаційний фон, а також виявлено їхні перспективи для подальшого розвитку.

Методи інтерполяції, такі як кригінг та інші геостатистичні моделі, у поєднанні з машинним навчанням, дозволяють досягти високої точності у прогнозах радіаційного забруднення. Використання штучного інтелекту та алгоритмів обробки Big Data роблять можливим автоматичне виявлення закономірностей у даних, що дозволяє точно прогнозувати радіаційний фон навіть у місцях, де відсутні безпосередні вимірювання.

2 ВИБІР СУЧАСНИХ ПІДХОДІВ ДО ОБРОБКИ ТА АНАЛІЗУ ДАНИХ

2.1 Вибір рішень для розв’язання поставленого завдання

З розвитком інформаційних технологій інформаційні системи стали невіддільною частиною технічної інфраструктури більшості організацій. Використання сучасних інформаційних систем дозволяє значно покращити організаційні процеси. Машинне навчання продемонструвало свою ефективність завдяки здатності швидко та точно обробляти дані.

Сучасний ринок пропонує велику кількість товарів та послуг. Для того, щоб організація залишалася конкурентоспроможною, необхідно адаптувати всі її внутрішні процеси. Оскільки інформаційні системи є частиною цих процесів, вони також можуть підлягати змінам.

З розвитком хмарних технологій з’явилися можливості для більш ефективного зберігання та обробки величезних обсягів даних, які генеруються автоматизованими системами моніторингу радіаційного фону. Такі платформи дозволяють зберігати дані в реальному часі, проводити їх аналіз і прогнозування майбутніх загроз.

Інструменти, такі як Kaggle, стали популярними для аналізу даних, пов’язаних з екологічними проблемами, в тому числі радіаційним забрудненням. Вони дозволяють фахівцям з обробки даних об’єднувати свої зусилля для вирішення глобальних проблем.

Основним елементом інформаційної системи є інформація, тому важливо розуміти, як з нею працювати для отримання корисних результатів. Інформаційна технологія є програмним продуктом, тому її спосіб обробки інформації буде специфічним. Потрібно визначити, яку інформацію і як обробляти. Для цього існує ряд критеріїв до інформації:

- Однорідність: Інформація має бути представлена в єдиному форматі. Це може бути текстова, числова або змішана інформація, залежно від завдання.

- Обсяг даних: Великий обсяг даних може містити багато шуму, що ускладнює обробку.
- Взаємозв'язок даних: Необхідно розуміти, як дані взаємопов'язані, оскільки відсутність зв'язків знижує ефективність обробки.
- Корисність: Інформація повинна мати практичну цінність.

Сучасні методи аналізу даних включають [11]:

- штучні нейронні мережі (ШНМ);
- факторний аналіз;
- регресійний аналіз;
- тестування панельних даних;
- кореляційний аналіз;
- семантичний аналіз;
- Data Mining.

Вибір методу аналізу залежить від основного завдання програмного продукту. Наприклад, для фінансових інформаційних систем, що працюють з числовими даними, підходять методи факторного аналізу, регресійного аналізу та аналізу панельних даних. Інформація, подана на вхід системи, може бути в змішаному або відфільтрованому вигляді. Змішані дані приводяться до однорідного вигляду за допомогою Data Mining або інших методів автоматичної обробки, таких як штучні нейронні мережі.

Для аналізу текстової інформації застосовуються методи семантичного аналізу, які є складними математичними завданнями і часто використовуються в системах зі штучним інтелектом. Ці методи ефективні у сфері торгівлі, маркетингу та журналістики.

Отже, вибір методу аналізу даних в інформаційних технологіях залежить від мети чи завдання, яке буде виконуватися програмним продуктом. Після визначення завдання слід встановити критерії до вхідної інформації та привести її до відповідного вигляду. Лише після цього можна переходити до вибору методу аналізу даних.

2.1.1 ArcGIS Online

Для вирішення поставленого метою завдання магістерської кваліфікаційної роботи було обрано інтернет-середовище ArcGIS Online, яке використовується для реалізації функцій картографування та аналізу та дозволяє створювати власний картографічний набір даних і виконувати його аналіз.

ArcGIS є повнофункціональною системою, що дозволяє збирати, управляти, аналізувати, обмінюватися та розподіляти географічну інформацію. Як світовий лідер серед платформ для побудови та використання геоінформаційних систем (ГІС), ArcGIS використовується в різних галузях, включаючи державне управління, бізнес, науку, освіту та ЗМІ. Платформа ArcGIS дозволяє публікувати географічну інформацію для доступу та використання будь-якими користувачами. Система доступна в будь-яких цифрових середовищах – сучасних браузерів, мобільних та десктопних додатках тощо.

ArcGIS Online містить інтерактивні карти та 3D макети, що дозволяє переглядати, вивчати та аналізувати географічні дані (рис. 2.1). Використовуючи ArcGIS Living Atlas of the World, динамічну колекцію карт, сцен, шарів даних, зображень, аналітику та програми зі спільноти ArcGIS, а також інструменти аналізу, які містить Map Viewer Classic, можна знаходити нові закономірності, визначати відповідні місця, збагачувати дані, визначати найближчі об'єкти та підсумовувати дані.

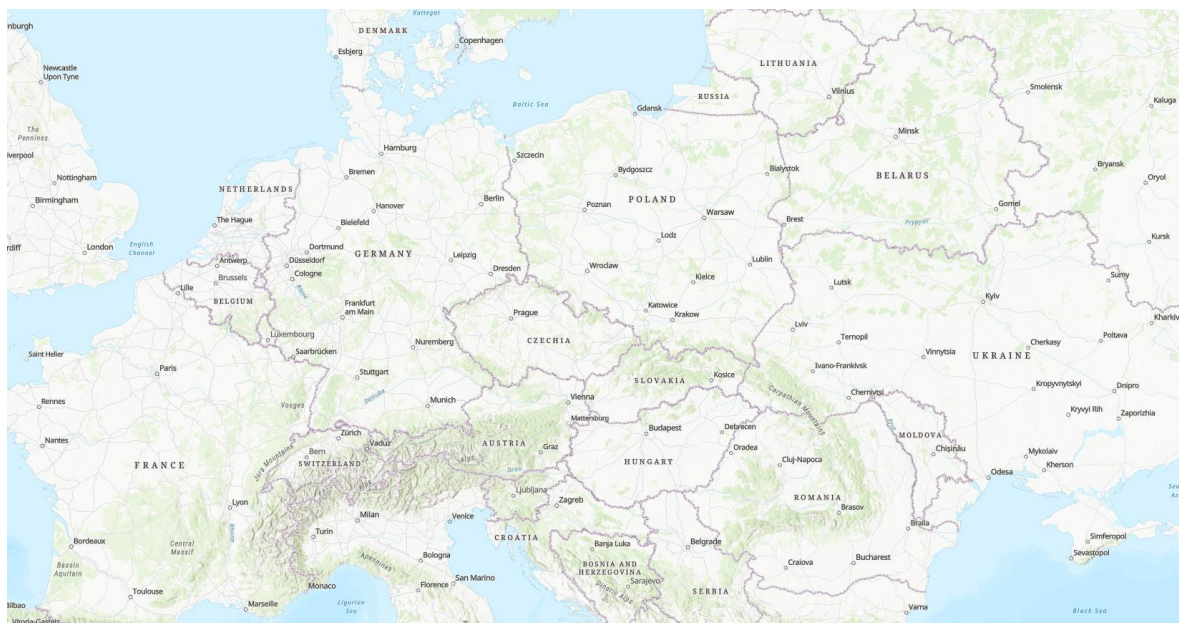


Рисунок 2.1 – Середовище ArcGIS online

Модуль Geostatistical Analyst надає потужні інструменти для інтерполяції даних та статистичного аналізу модельних поверхонь. Коли дані представлені не повністю або містять помилки, ArcGIS Geostatistical Analyst забезпечує ймовірнісну основу для кількісної оцінки невизначеностей та виключення помилок з прогнозу. Інтерполяція поверхонь виконується за допомогою таких методів [15]:

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

Було створено веб-сервіси для автоматизації процесу оцінювання параметрів джерел понаднормативних викидів на основі нечіткої бази знань з візуалізацією результатів у Google Maps, використовуючи запропонований метод та інформаційну технологію. 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 Python як основний інструмент аналізу даних

Мова програмування Python відіграє ключову роль у виконанні дослідження що було описано в даній роботі. Завдяки своїй гнучкості, багатофункціональності та великій кількості спеціалізованих бібліотек, Python дозволив забезпечити ефективну обробку, аналіз та візуалізацію даних. Конкретно у цьому дослідженні Python використовувався для роботи з великими обсягами даних, застосування методів інтерполяції, створення геостатистичних моделей та побудови графічних візуалізацій результатів [13].

Протягом виконання роботи було використано такі бібліотеки:

- Pandas – основна бібліотека для роботи з табличними даними. У цьому дослідженні вона використовувалась для завантаження, очищення та структуризації даних взятих з датасету. За її допомоги виконувалася нормалізація наборів даних, заповнення пропусків та перетворення форматів, що забезпечило підготовку даних до аналізу.
- NumPy – бібліотека, роль якої відобразилась у виконанні числових операцій із великими наборами даних. У дослідженні вона використовувалась для виконання математичних обчислень, зокрема інтерполяції, що стало основою побудови карт радіаційного забруднення.
- Matplotlib та Seaborn використовувались для візуалізації результатів дослідження. Matplotlib забезпечував базову побудову графіків і карт, тоді як Seaborn дозволяв створювати більш складні графічні представлення залежностей між даними, наприклад, теплові карти або гістограми розподілу.
- SciPy використовувався для реалізації інтерполяційних методів, таких як метод зворотного зважування відстаней (IDW). Ця бібліотека є важливою для геостатистичних обчислень, які дозволяють інтерполювати показники радіаційного фону у місцях, де відсутні вимірювання.
- SciKit-GStat — спеціалізована бібліотека для геостатистичного аналізу, яка використовувалась для побудови варіограм та моделювання просторових залежностей у даних. Це дозволило точніше визначити закономірності

поширення радіаційного забруднення.

- GeoPandas розширює функціонал Pandas для роботи з геопросторовими даними. У цьому дослідженні вона використовувалася для обробки географічної інформації, створення карт радіаційного забруднення та інтеграції даних з GIS.
- Бібліотека Scikit-learn забезпечила використання алгоритмів машинного навчання для аналізу та кластеризації даних. Це дозволило ефективно визначати групи станцій із подібними рівнями забруднення та моделювати динаміку поширення радіації.
- Folium використовувався для створення інтерактивних карт радіаційного забруднення. Це дозволило візуалізувати дані у вигляді динамічних картографічних моделей, які легко інтегруються з іншими інструментами.

2.1.3 Kaggle як платформа для виконання обчислень

Kaggle є відомою платформою для здійснення досліджень у галузях Data Science та machine learning, широко визнаною завдяки своїй великій спільноті та ресурсам. Він пропонує середовище для спільної роботи, де інженери з обробки даних, фахівці з машинного навчання та дослідники можуть працювати над проектами разом, брати участь у конкурсах, ділитися наборами даних і ноутбуками. Екосистема Kaggle включає ряд інструментів і послуг, призначених для оптимізації робочого процесу аналізу даних і розробки моделей.

Серед ключових переваг сервісу можна виокремити:

- Kaggle регулярно проводить численні змагання, де учасники можуть вирішувати реальні задачі та вигравати призи. Це дає мотивацію для навчання та привід для потенційного визнання.
- Величезне сховище загальнодоступних наборів даних, доступних для аналізу та побудови моделей.
- Інтерактивні блокноти (ноутбуки) Jupyter, які дозволяють кодувати, аналізувати та ділитися результатами в межах спільноти.
- Активна спільнота ентузіастів аналізу даних, які активно діляться ідеями, кодом та надають підтримку за допомогою спеціалізованих

форумів та обговорень.

- Безліч безкоштовних курсів з Data Science та машинного навчання, які допомагають користувачам постійно покращувати свої навички.

2.2 Дослідження інтерполяції даних як явища

Інтерполяція даних є однією з ключових методик у сфері аналізу даних, статистики та обчислювальної математики. Вона дозволяє на основі відомих значень функції у певних точках відновлювати значення цієї функції в інших точках, де вони невідомі. Це робить інтерполяцію незамінним інструментом у багатьох галузях науки та техніки, включаючи фізику, геологію, економіку, екологію та медицину.

Дослідження інтерполяції як явища містить в собі вивчення різних методів і алгоритмів, таких як лінійна інтерполяція, поліноміальна інтерполяція, сплайнові методи та методи на основі радіальних базисних функцій. Кожен з цих методів має свої переваги та недоліки, що визначає їх застосовність у конкретних ситуаціях.

Основна мета дослідження інтерполяції полягає у знаходженні найбільш точних і ефективних способів відновлення даних, що дозволяє отримувати правдиві результати навіть у випадках, коли обсяги доступної інформації є обмеженими. Це, своєю чергою, сприяє підвищенню точності обчислень, оптимізації процесів та прийняття обґрунтованих рішень на основі аналізу даних.

У сучасному світі, де обсяг даних стрімко зростає, а вимоги до їх обробки та аналізу стають все більш суворими, інтерполяція даних займає важливе місце у наукових дослідженнях та практичних застосуваннях. Вивчення цього явища дозволяє розширити наші можливості у роботі з даними, розробляти нові методики та вдосконалювати існуючі алгоритми, що сприяє прогресу у різних галузях знань та технологій.

2.2.1 Інтерполяція даних за допомогою методу Kriging

Кригінг — це вдосконалений метод інтерполяції, який використовується в інженерії даних для прогнозування значень у місцях без вибірки шляхом використання лінійної комбінації спостережень із найближчих точок вибірки. Цей метод оцінює вплив кожного спостереження на основі трьох основних факторів:

- 1) географічної близькості спостереження до місця без вибірки;
- 2) просторового розташування всіх спостережень (наприклад, групування в районах із надмірною вибіркою);
- 3) шаблону просторова кореляція в даних. Ефективність моделей кригінгу залежить від наявності просторово корельованих даних.

Кригінг пропонує кілька переваг порівняно з традиційними методами інтерполяції, такими як зворотне зважування відстані (IDW) або інтерполяція найближчого сусіда:

- Кількісна оцінка невизначеності: кригінг надає міру невизначеності (дисперсія кригінга), пов'язану з його прогнозами.
- Просторова анізотропія: враховує зв'язки, що залежать від напрямку, підвищуючи точність у просторово анізотропних середовищах.
- Ваги на основі даних: ваги призначаються спостереженням на основі просторової кореляції даних, а не довільних припущень.
- Гнучкість діапазону: передбачення кригінгу не обмежуються діапазоном спостережуваних даних, які використовуються для інтерполяції.
- Інтеграція підтримки: це дозволяє інтегрувати дані, виміряні на різних просторових опорах, і полегшує зміну операцій підтримки, таких як зменшення або збільшення масштабу.

Ці особливості роблять кригінг потужним інструментом у розробці даних, особливо в галузях, що вимагають точних просторових прогнозів і оцінки невизначеності.

Кригінг-метод 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 [6]. Сфера його застосування значно розширилася, щоб охопити сьогодні більшість галузей геонауки (наприклад, геологію, геохімію, геогідрологію, ґрунтознавство тощо) і широкий спектр дисциплін, які займаються аналізом просторово-часових даних як от океанографія, гідрогеологія, дистанційне зондування, сільського господарства, географії, ГІС та природокористування.

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 .

Дві двовимірні конфігурації даних з різними положеннями для розташування u_3 . В обох випадках передбачення в місці u_0 виконується з використанням звичайного кригінга та п'яти даних у розташуваннях $u_1 - u_5$.

Радіус кола з центром u_0 відповідає діапазону кореляції [14] (рис. 2.2).

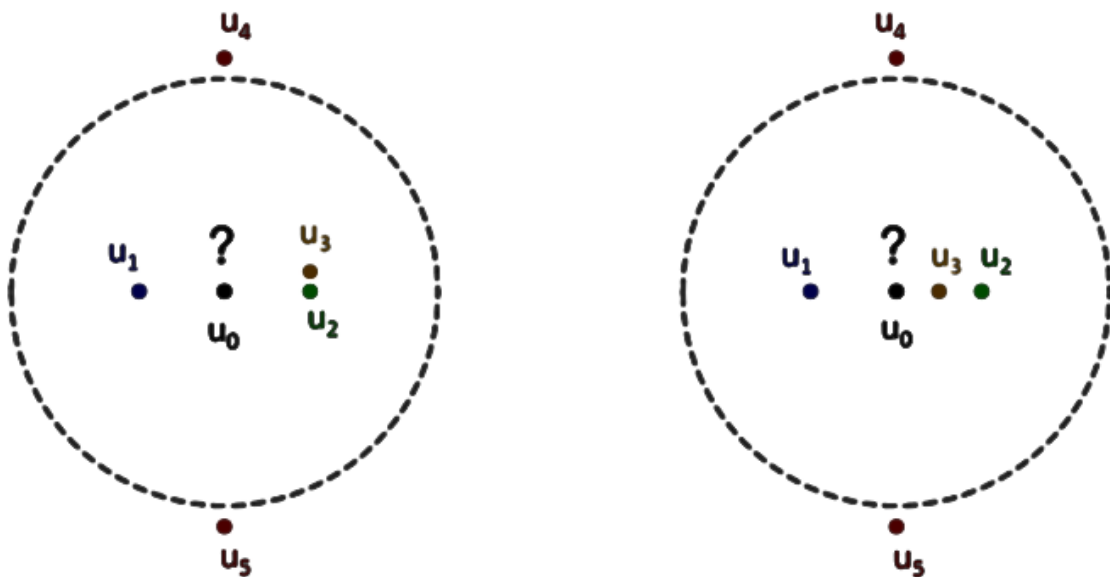


Рисунок 2.2 – Спрощений приклад кригінг-інтерполяції

2.2.2 Інструменти інтерполяції

Інструменти інтерполяції поверхні дозволяють створювати безперервні (або прогнозовані) поверхні на основі вибірових точкових значень. Вимірювання кожної точки в зоні дослідження, щоб визначити висоту,

концентрацію або величину певного явища, зазвичай є складним або дорогим процесом. Натомість, явище можна вимірювати в стратегічно розташованих вибіркових точках, а прогнозовані значення призначати всім іншим місцям. Вхідні точки можуть бути вибрані випадковим чином, регулярним чином або за певною вибірковою схемою.

Неперервне представлення поверхні у вигляді растрових даних відображає певну міру, таку як висота, концентрація або величина (наприклад, висота, кислотність або рівень шуму). Інструменти інтерполяції поверхні роблять прогнози на основі вибіркових вимірювань для всіх місць у вихідному наборі растрових даних, незалежно від того, чи було проведено вимірювання в конкретному місці чи ні.

Існує багато методів для інтерполяції значень у кожному місці; кожен метод називається моделлю. Кожна модель має свої припущення щодо даних, і певні моделі можуть краще підходити для конкретних даних — наприклад, одна модель може краще враховувати місцеві варіації, ніж інша. Кожна модель використовує різні розрахунки для створення прогнозів.

Інструменти інтерполяції можна загалом поділити на детерміновані та геостатистичні методи.

Детерміновані методи інтерполяції призначають значення точкам на основі навколишніх вимірних значень і застосовують задані математичні формули для визначення гладкості результуючої поверхні. До таких методів належать зважування зворотної відстані (IDW), Natural Neighbor, Trend та Spline.

Геостатистичні методи базуються на статистичних моделях, які враховують автокореляцію (статистичний зв'язок між вимірними точками). Завдяки цьому, геостатистичні методи не лише здатні створювати поверхні прогнозів, але й надають міру впевненості або точності цих прогнозів. Кригінг є одним з геостатистичних методів інтерполяції.

Інші інструменти інтерполяції, такі як Topo to Raster і Topo to Raster by File, використовують спеціалізовані методи інтерполяції для створення безперервних

поверхонь із контурних ліній, які мають властивості, сприятливі для гідрологічного аналізу.

2.2.3 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.

Інструмент IDW (Inverse Distance Weighted) використовує метод інтерполяції, що оцінює значення комірки шляхом усереднення значень вибірових точок даних в околиці кожної комірки обробки. Чим ближче точка до центру комірки, тим більший вплив або вона має на усереднення (рис. 2.3).

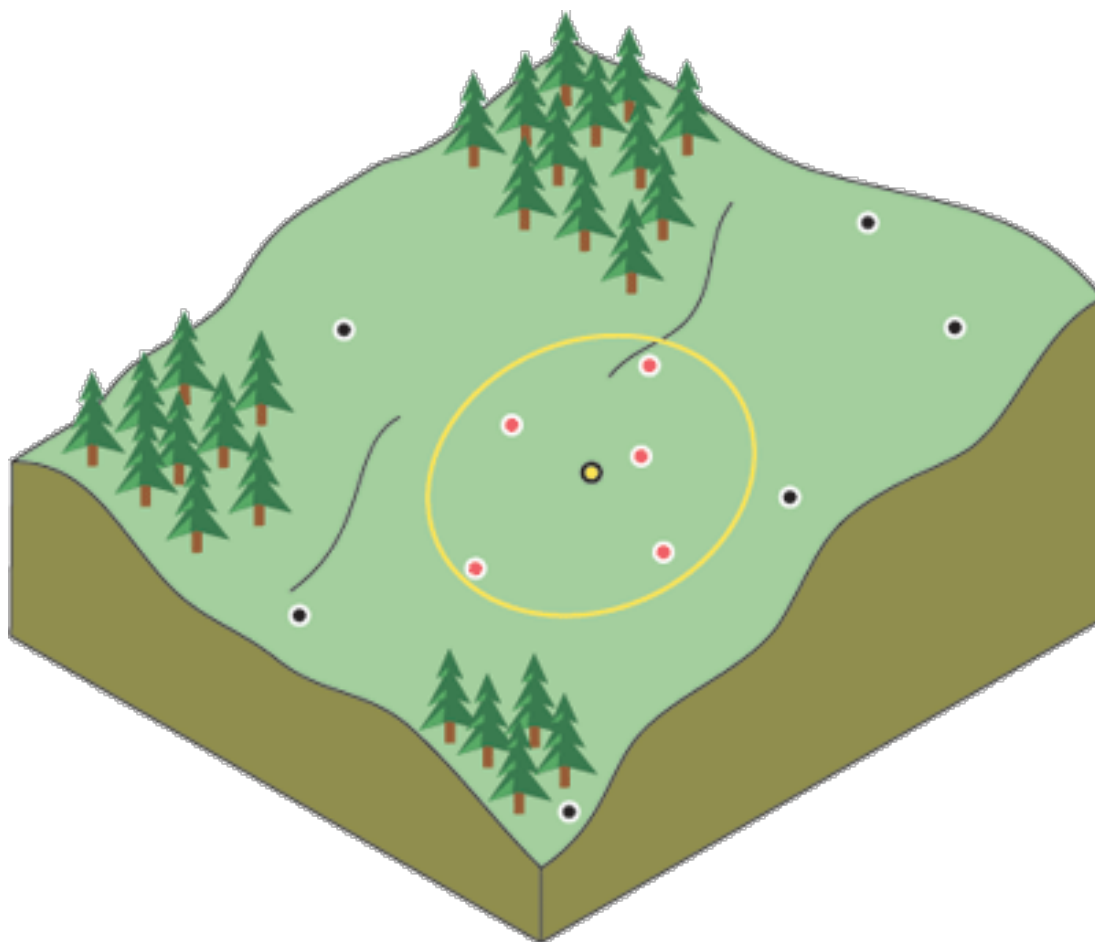


Рисунок 2.3 – Метод інтерполяції IDW

Кригінг – це вдосконалена геостатистична процедура, яка генерує оцінку поверхні з розсіяного набору точок із значеннями z (рис. 2.4). 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].

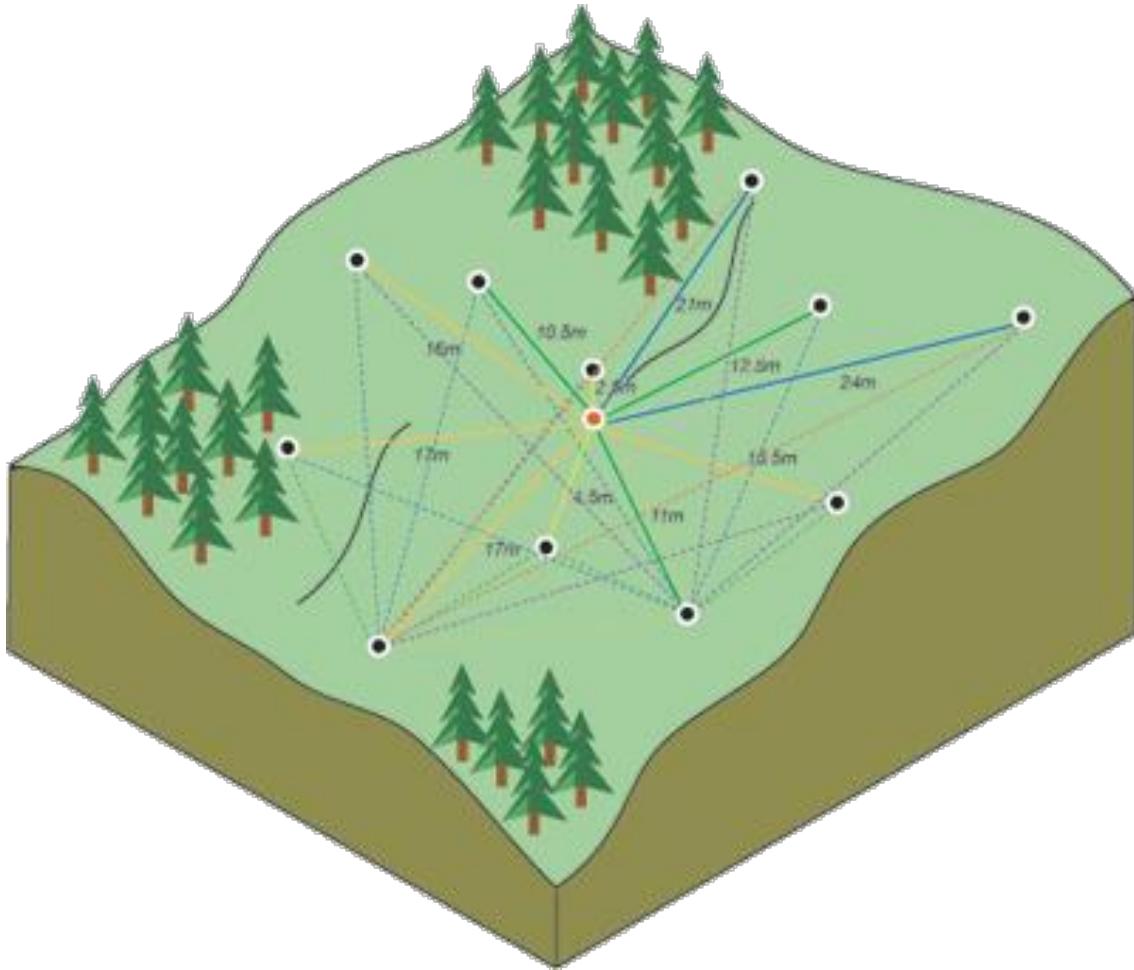


Рисунок 2.4 – Метод інтерполяції Кригінг

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). Вона також відома як інтерполяція Сібсона або інтерполяція, що викрадає область (рис 2.5).

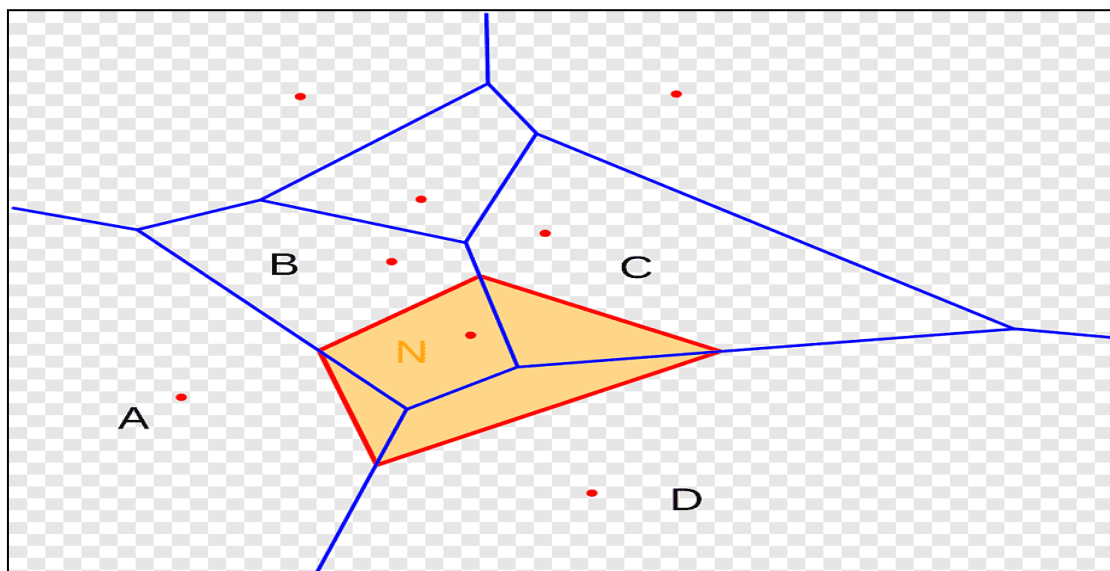


Рисунок 2.5 – Інтерполяція Natural Neighbor

Інструмент Сплайн використовує метод інтерполяції, який оцінює значення за допомогою математичної функції, яка мінімізує загальну кривину поверхні, що приводить до гладкої поверхні, яка проходить точно через вхідні точки (рис. 2.6) [14].

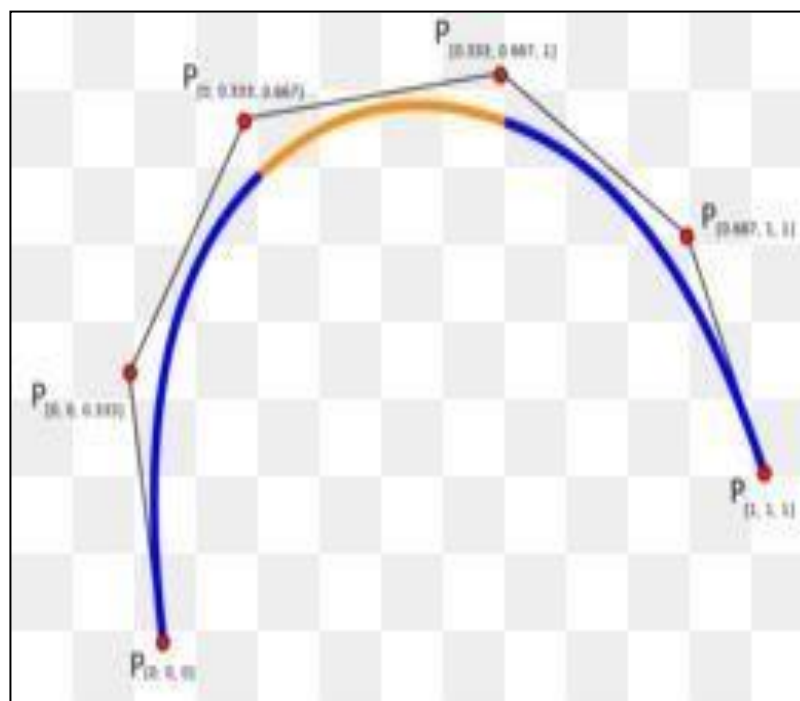


Рисунок 2.6. Сплайн-інтерполяція

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, задовані як у вхідних бар'єрах, так і в даних точки введення (рис. 2.7).

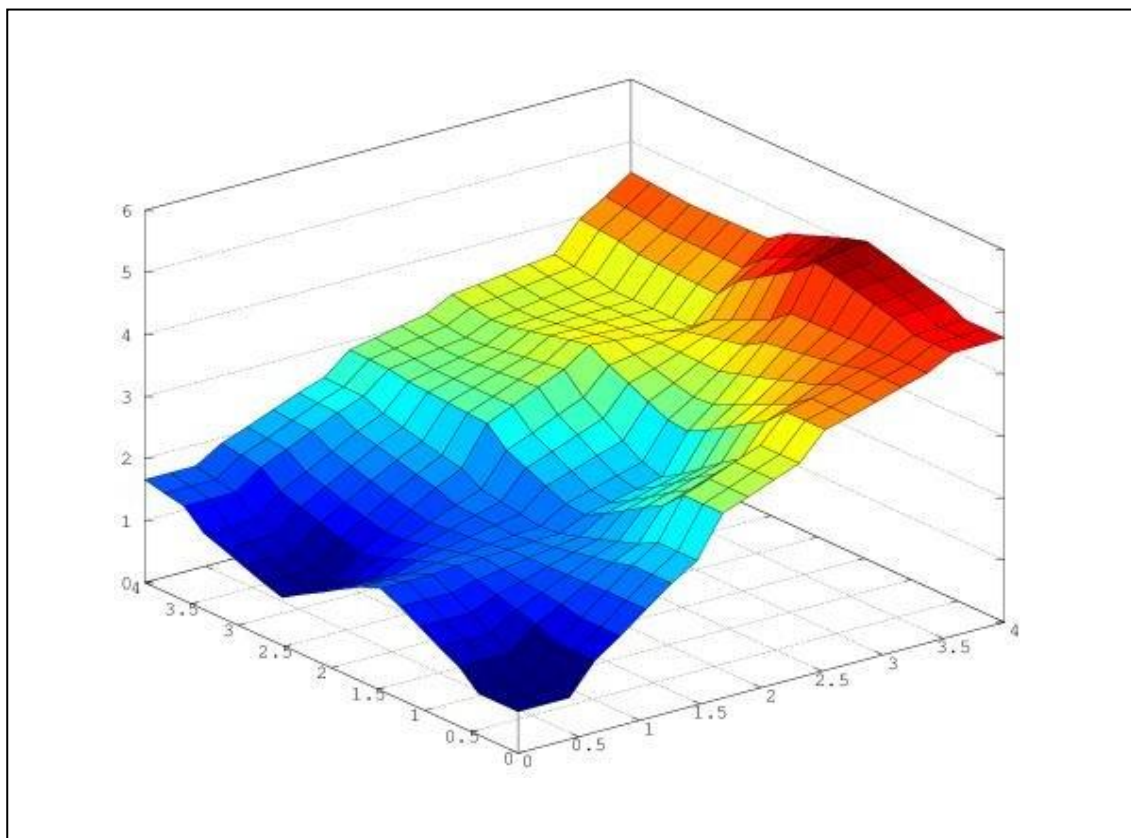


Рисунок 2.7 – Інструмент «Сплайн із бар'єрами»

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 поточкові мережі від вхідних даних контуру (рис. 2.8) [8].

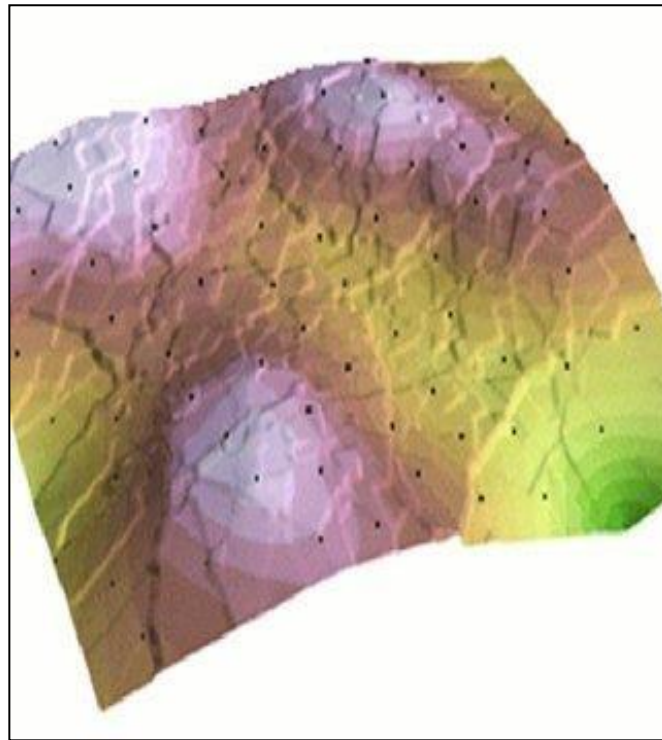


Рисунок 2.8 – Інструменти Tpo to Raster

Використовуваний алгоритм базується на алгоритмі ANUDEM, розробленому Хатчинсоном та іншими в Австралійському національному університеті [20].

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. 2.9).

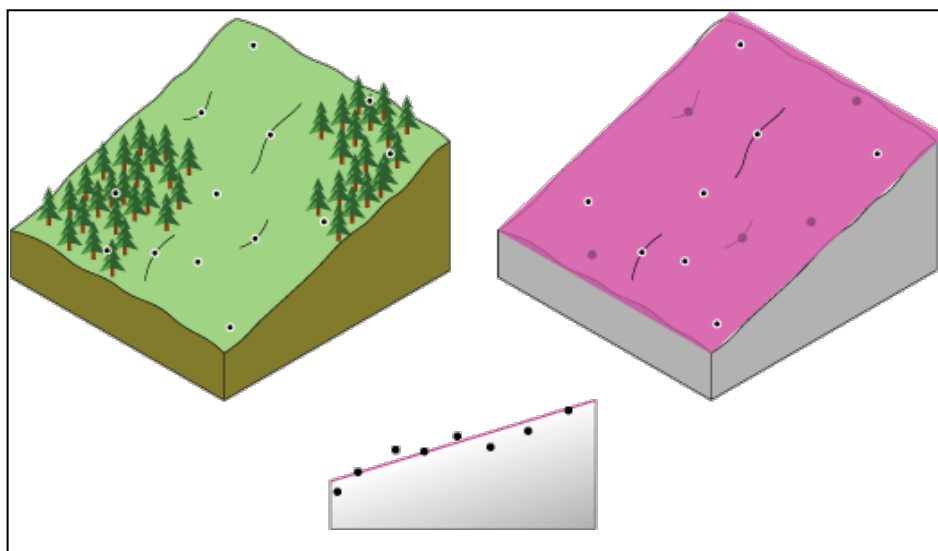


Рисунок 2.9 – Тренд-інтерполяція

2.3 Висновки

У даному розділі було проведено детальний аналіз технічної складової та вибору оптимальних рішень для вирішення поставленого завдання з використанням інформаційних технологій - з розвитком сучасних інформаційних технологій, використання сучасних інформаційних систем, таких як ArcGIS Online та Kaggle, стало ключовим для покращення організаційних процесів та аналізу даних.

Дослідження інтерполяції даних, зокрема за допомогою методів Kriging та інших геостатистичних методів, виявилось важливим етапом у моделюванні просторового розподілу радіаційного забруднення. Вибір оптимального методу аналізу даних у інформаційних технологіях визначається метою та завданням програмного продукту, що підкреслює важливість правильного підходу до обробки та аналізу інформації.

Загальний висновок полягає в тому, що використання сучасних методів аналізу даних, таких як штучні нейронні мережі та факторний аналіз, разом з інструментами геостатистичного аналізу та ГІС, дозволяє ефективно вирішувати завдання та досягати поставлених цілей у дослідженні радіаційного забруднення.

3 РОЗРОБЛЕННЯ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ

3.1 Підготовка вхідних даних

На початковому етапі дослідження важливо забезпечити належну підготовку вхідних даних для подальшого аналізу та візуалізації. Оскільки у дослідженні використовуються дані про радіаційне забруднення, необхідно провести нормалізацію даних, видалити пропуски та невірні значення, а також стандартизувати одиниці виміру.

Зібрані дані можуть походити з різних джерел, таких як наземні станції, мобільні моніторингові системи або супутникові платформи. Кожен із цих джерел може мати свої специфічні формати, що потребують конвертації та уніфікації для їхнього подальшого аналізу. Для цього використовуються спеціальні інструменти обробки даних, такі як Python із бібліотеками Pandas та NumPy, що дозволяють виконати очищення та нормалізацію даних.

Обов'язковим етапом перед проведенням дослідження є безпосередньо підготовка даних до нормалізованої форми, тобто до такої яку можна без проблем використовувати для розрахунків та побудови графіків.

У якості вхідних даних у даному дослідженні використовується датасет Chernobyl Chemical Radiation / CSV / Country Data, що містить в собі дані про рівень радіаційного фону радіонуклідів Йод-131, Цезій-134 та Цезій-137, що вимірювались на території сучасної Європи після катастрофи на Чорнобильській АЕС.

Дані мають загальний вигляд що зображений на рисунку 3.1.


```
In [4]: df = pd.read_csv('/kaggle/input/air-concentration-for-the-chernobyl-disaster/data.csv', header=0,
sep=',')
df
```

Out[4]:

	PAYS	Code	Ville	X	Y	Date	End_of_sampling	Duration(h.min)	L_131_(Bq/m3)	Cs_134_(Bq/m3)
0	SE	1	RISOE	12.07	55.7	86/04/27	24:00:00	24.0	1	0
1	SE	1	RISOE	12.07	55.7	86/04/28	24:00:00	24.0	0.0046	0.00054
2	SE	1	RISOE	12.07	55.7	86/04/29	12:00	12.0	0.0147	0.0043
3	SE	1	RISOE	12.07	55.7	86/04/29	24:00:00	12.0	0.00061	0
4	SE	1	RISOE	12.07	55.7	86/04/30	24:00:00	24.0	0.00075	0.0001
...	...	...	...	...	...	...	...	...	...	...
2046	HU	24	BUDAPEST	19.10	47.5	86/05/07	08:00	24.0	3.7	0.58
2047	HU	24	BUDAPEST	19.10	47.5	86/05/08	08:00	24.0	2.2	0.33
2048	HU	24	BUDAPEST	19.10	47.5	86/05/09	08:00	24.0	0.75	0.14
2049	HU	24	BUDAPEST	19.10	47.5	86/05/10	08:00	24.0	0.054	0.006
2050	HU	24	BUDAPEST	19.10	47.5	86/05/11	08:00	24.0	0.051	0.006

Рисунок 3.1 – Загальний вигляд даних

Для зручності роботи з даними обмежуємо область ознак лише до необхідних значень. Отримуємо наступний вигляд вхідних даних (рис. 3.2).

Для кращого розуміння характеристики наявних даних будемо картографічне відображення за допомогою бібліотеки Folium (рис. 3.3).

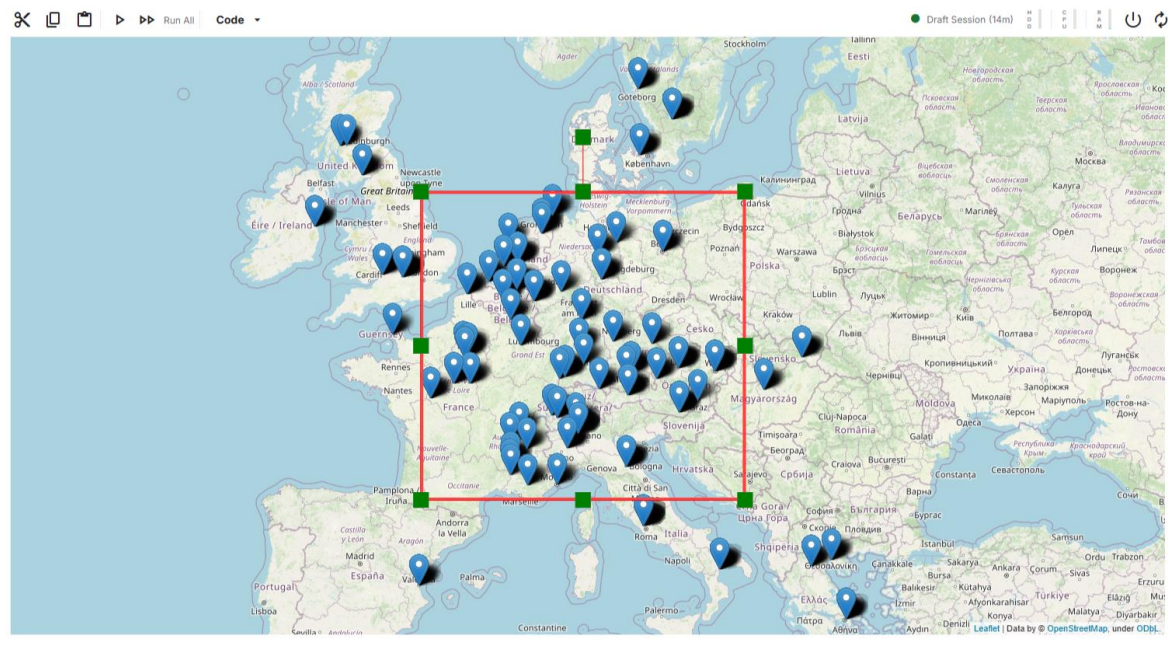


Рисунок 3.3 – Картографічне відображення вхідних даних

Виконавши базовий аналіз даних, приймаємо рішення що для здійснення геостатистичного аналізу нам необхідно відфільтрувати дані та використовувати

лише дані країн, які мають найбільшу щільність станцій виміру радіаційного фону. Відповідно, приймаємо рішення виконувати аналіз враховуючи лише дані таких країн як Бельгія, Швейцарія, Німеччина, Франція, Нідерланди, Австрія та Чехія (рис. 3.4) для забезпечення найкращої точності аналізу.

```
countries = ['Belgium', 'Switzerland', 'Germany', 'France', 'Netherlands', 'Austria', 'Czech' ]
df = df.loc[df['PAYS'].isin(countries)]
```

Рисунок 3.4 – Вибірка країн для дослідження

Часто у великих наборах даних зустрічаються записи з відсутніми значеннями. У контексті радіаційного моніторингу такі пропуски можуть бути викликані відсутністю показників із певних станцій або технічними збоями у передачі даних. Саме такий випадок і відбувся у цьому дослідженні, і тому було прийнято рішення використати інтерполяцію даних, що дозволить відновити відсутні значення на основі доступних даних.

3.2 Кластеризація та інтерполяція даних

Після обробки та нормалізації даних, необхідно провести їх кластеризацію. У контексті моніторингу радіаційного фону, кластеризація допомагає виявити групи станцій із подібними рівнями забруднення та створити моделі для передбачення подальших змін. Найбільш поширеним підходом є застосування методу K-means, який дозволяє розділити дані на певну кількість кластерів на основі показників радіаційного забруднення.

Метод кластеризації K-means дозволяє розділити набір даних на кілька кластерів, у яких кожна точка належить до найближчого центроїду. Це особливо корисно для визначення регіонів з подібними рівнями забруднення, що дозволяє побудувати зони впливу забруднення та зрозуміти, як радіоактивні речовини поширюються на різних територіях.

Проведемо додатковий аналіз щільності розташування отриманих даних за допомогою діаграми розсіювання. Для цього використовуємо бібліотеку

Matplotlib та отримуємо дані зображені на рисунку 3.5.

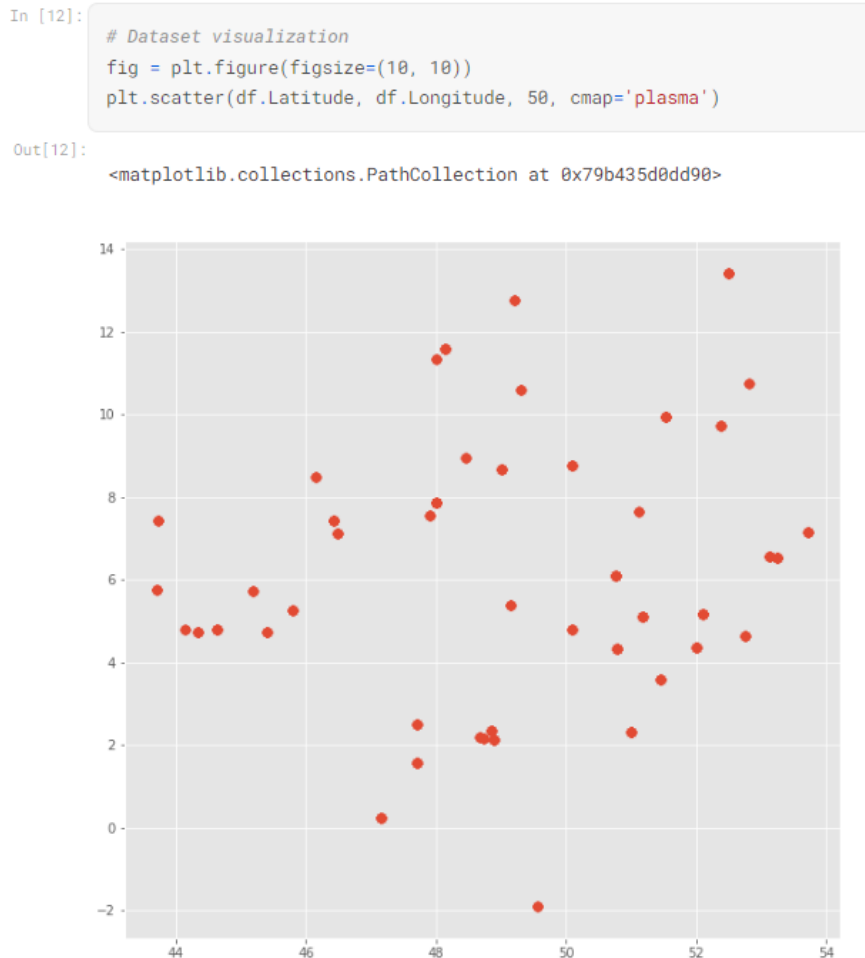


Рисунок 3.5 – Діаграма розсіювання даних

Помічаємо що дані розташовані достатньо щільно, проте внаслідок географічних особливостей забір даних у певних ділянках рельєфу не є можливим.

Для корекції даних приймаємо рішення здійснити кластеризацію та передбачення даних за допомогою алгоритму K-Nearest Neighbor.

Для цього встановлюємо оптимальну кількість кластерів розбиття за допомогою «ліктьового методу» враховуючи дані кожної з важливих ознак – рівню заміру нуклідів Йод-131, Цезій-134 та Цезій-137 та помічаємо що оптимальне значення – 4 (рис. 3.6).

```

Sum_of_squared_distances = []
K = range(2,8)
for num_clusters in K :
    kmeans = KMeans(n_clusters=num_clusters)
    kmeans.fit(df[["Iodine_131", "Caesium_134", "Caesium_137"]].notna())
    Sum_of_squared_distances.append(kmeans.inertia_)
plt.plot(K, Sum_of_squared_distances, "bx-")
plt.xlabel("Values of K")
plt.ylabel("Sum of squared distances/Inertia")
plt.title("Elbow Method For Optimal k")
plt.show()

```

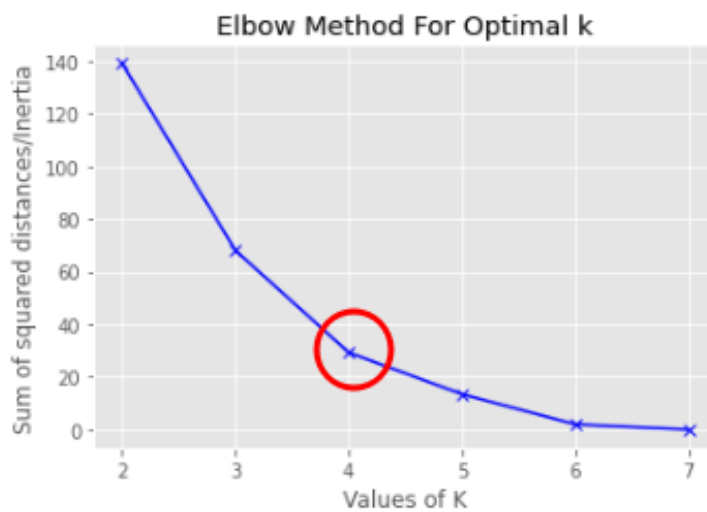


Рисунок 3.6 – "Ліктювий метод" визначення кількості кластерів

Виконавши кластеризацію розпочинаємо генерацію даних передбачених за допомогою K-nearest Neighbor алгоритму. Встановлюємо оптимальне значення 3 сусідів для кожного із вже існуючих записів та виконуємо передбачення. Отримуємо показник точності передбачення у 96.5% що свідчить про високий ступінь схожості даних до реальних замірів (рис 3.7).

```

# Predict on test data
y_pred = knn.predict(X_test)

# Calculate accuracy
accuracy = accuracy_score(y_test, y_pred)
print(f"Accuracy: {accuracy:.3f}")

# Predict clusters for new data
new_data = np.array([[4.00, 51.00, 0.15, 0.12, 0.25], [7.00, 47.00, 0.25, 0.18, 0.30]])
new_data_scaled = scaler.transform(new_data)
new_clusters = knn.predict(new_data_scaled)
print("Predicted clusters for new data:", new_clusters)

Accuracy: 0.965

```

Рисунок 3.7 – Передбачення даних за допомогою K-nearest Neighbor алгоритму

Встановлюємо мінімальний показник у 5% додаткових точок при генерації даних та отримуємо наступну діаграму розсіювання даних (рис. 3.8). Здійснюємо візуальний аналіз та робимо висновок що дані готові до виконання геостатистичного аналізу.

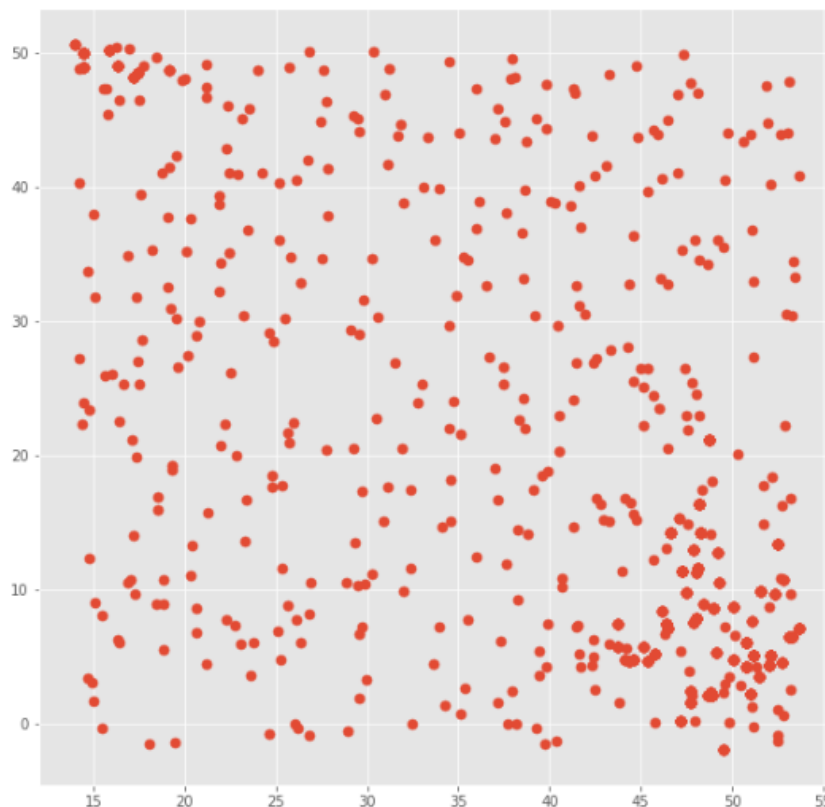


Рисунок 3.8 – Діаграма розсіювання даних після передбачення допоміжних даних

Результати машинного навчання можна використовувати для побудови довгострокових прогнозів, що дозволяє визначити потенційні небезпеки, а також розробити плани щодо мінімізації впливу радіації на довкілля.

3.3 Геостатистичний аналіз за допомогою SciKit-GStat

Геостатистичний аналіз є важливою частиною дослідження, оскільки він дозволяє виявити просторові закономірності та аномалії в даних про радіаційне забруднення. У нашому дослідженні використовуються інструменти бібліотеки SciKit-GStat, яка пропонує широкий набір функцій для побудови варіограм і моделювання просторових залежностей.

На основі аналізу даних з отриманої діаграми розсіювання будемо сферичні варіограми для основних досліджуваних ознак - Йод-131 (рис. 3.9), Цезій-134 (рис. 3.10) та Цезій-137 (рис. 3.11).

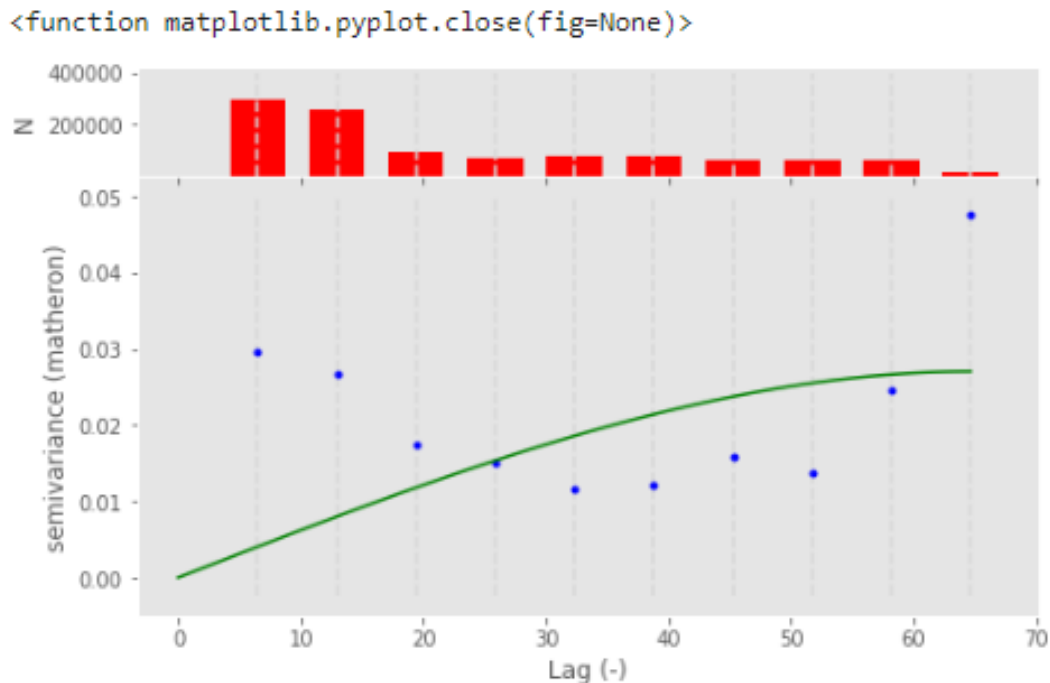


Рисунок 3.9 – Сферична варіограма по показнику Йод-131

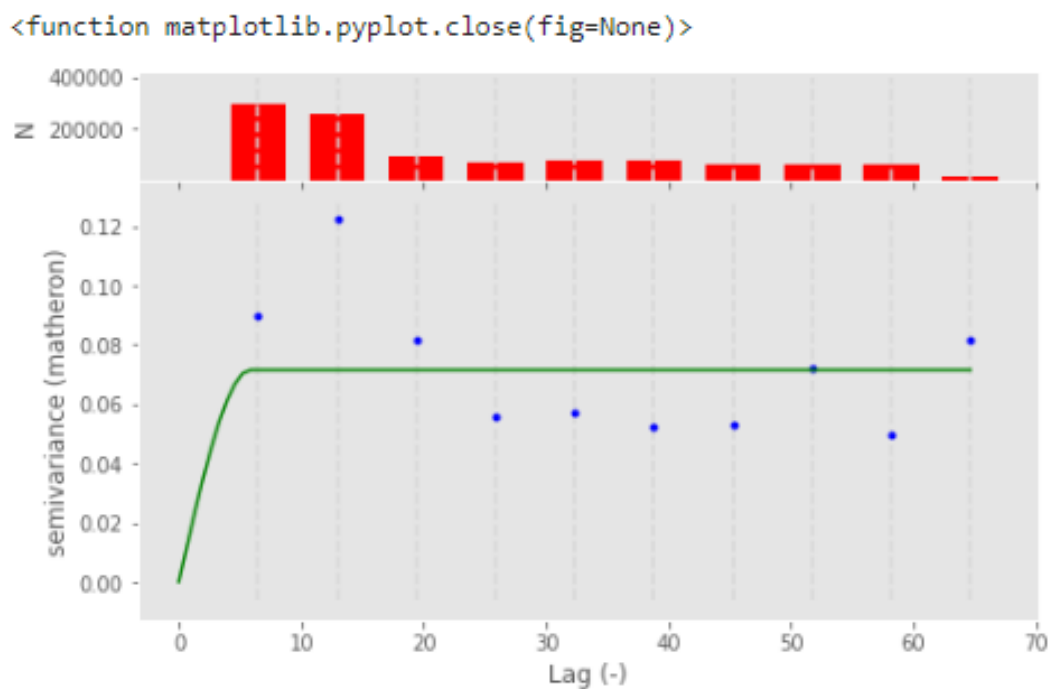


Рисунок 3.10 – Сферична варіограма по показнику Цезій-134

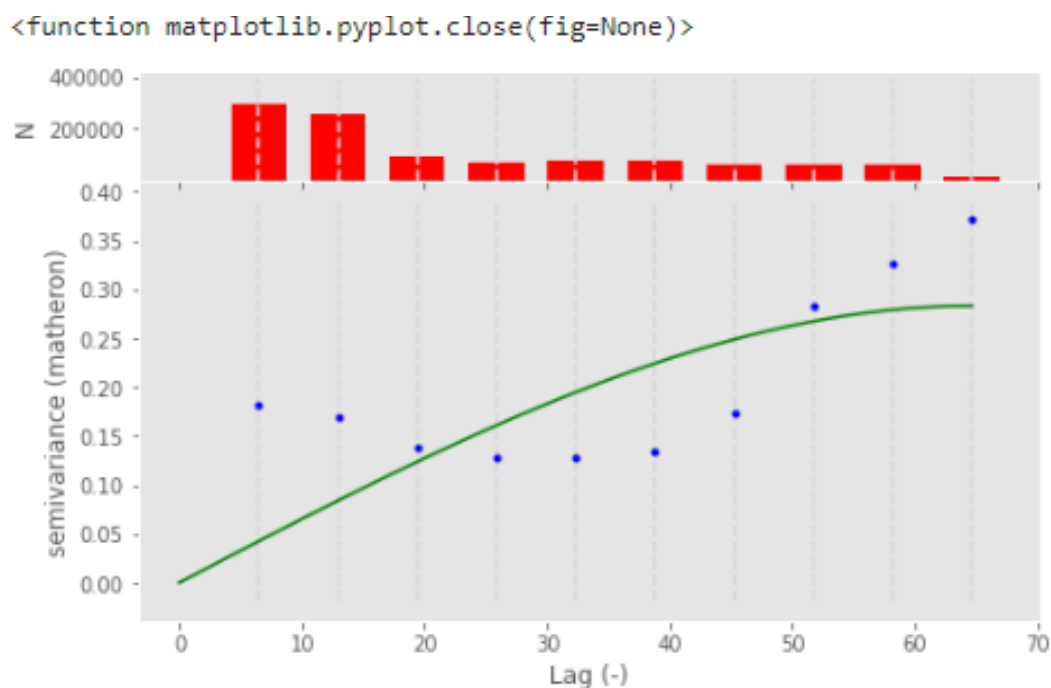


Рисунок 3.11 – Сферична варіограма по показнику Цезій-137

На даних рисунках можливо побачити на скільки далеко чи близько один від одного знаходяться станції відносно своєї широти.

На основі побудованих варіограм вимальовується модель просторових залежностей, яка дозволяє зробити більш точне визначення рівня радіації у невідомих точках.

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 станцій там, де помітні різкі зміни показників заміру радіаційного забруднення, а по-друге, там – де станій замало (зазвичай у долинах морів/річок та у гористій місцевості).

В якості експерименту порівнюємо різні моделі побудови варіограм, проте помічаємо що на наших даних вони всі дають приблизно однакові результати (рис. 3.12).

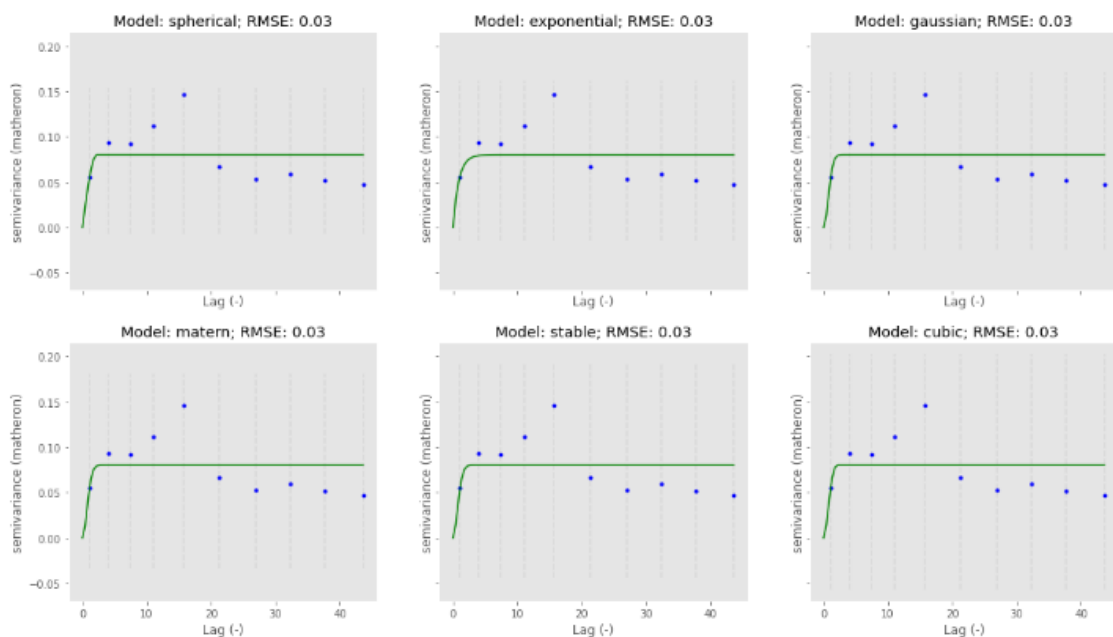


Рисунок 3.12 – Порівняння основних моделей побудови варіограм

Приступаємо безпосередньо до візуалізації даних. На попередньо опрацьованих даних отримуємо показники інтерполяції (рис. 3.13) та Kriging Error (рис. 3.14).

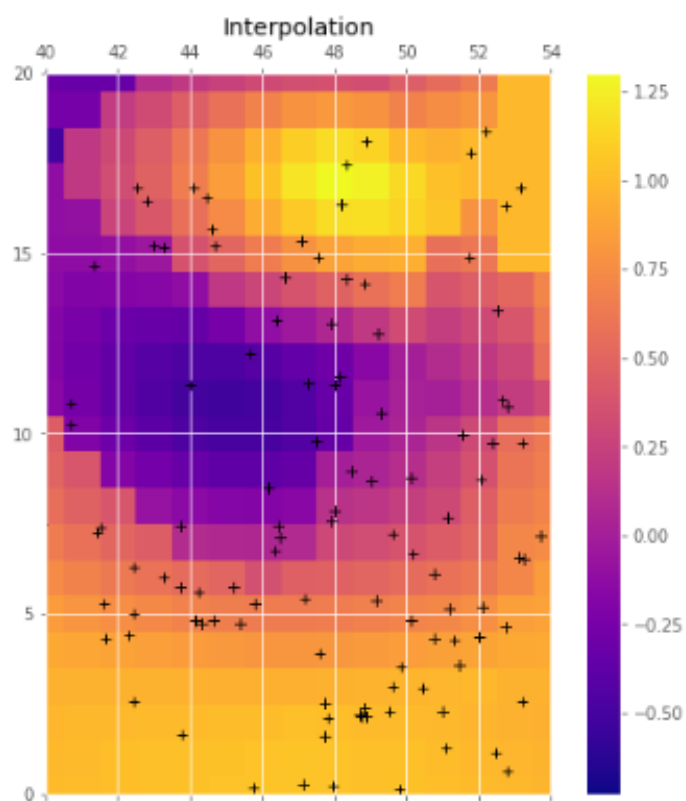


Рисунок 3.13 – Графік інтерполяції

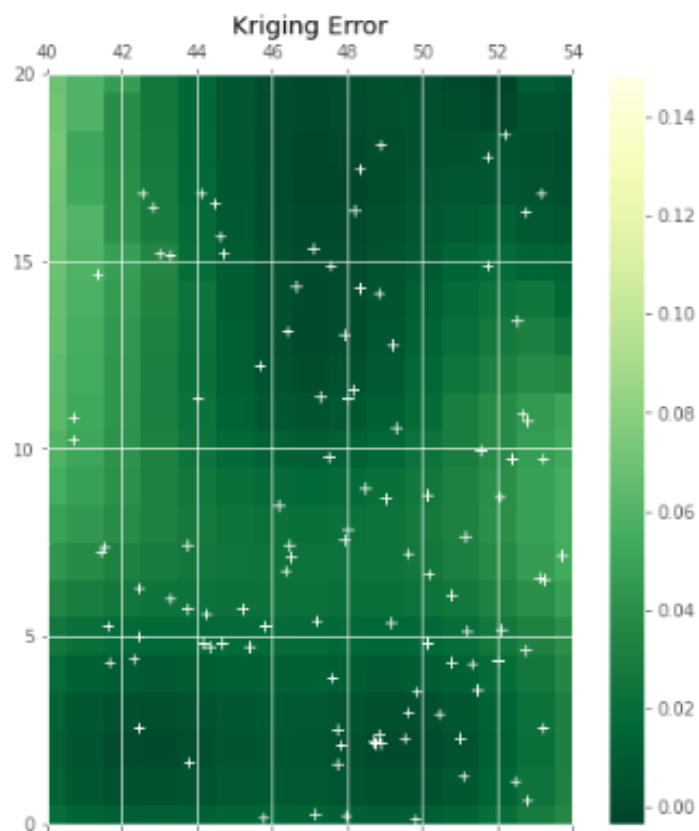


Рисунок 3.14 – Графік Kriging Error

Важливо що для підвищення ефективності аналізу необхідно задавати значення не випадковим чином, а на основі певної математичної функції, додаючи шум, щоб метод міг знаходити чинні закономірності.

Робимо геостатистичний аналіз на відповідних даних та отримуємо очікувані результати на рисунку 3.15.

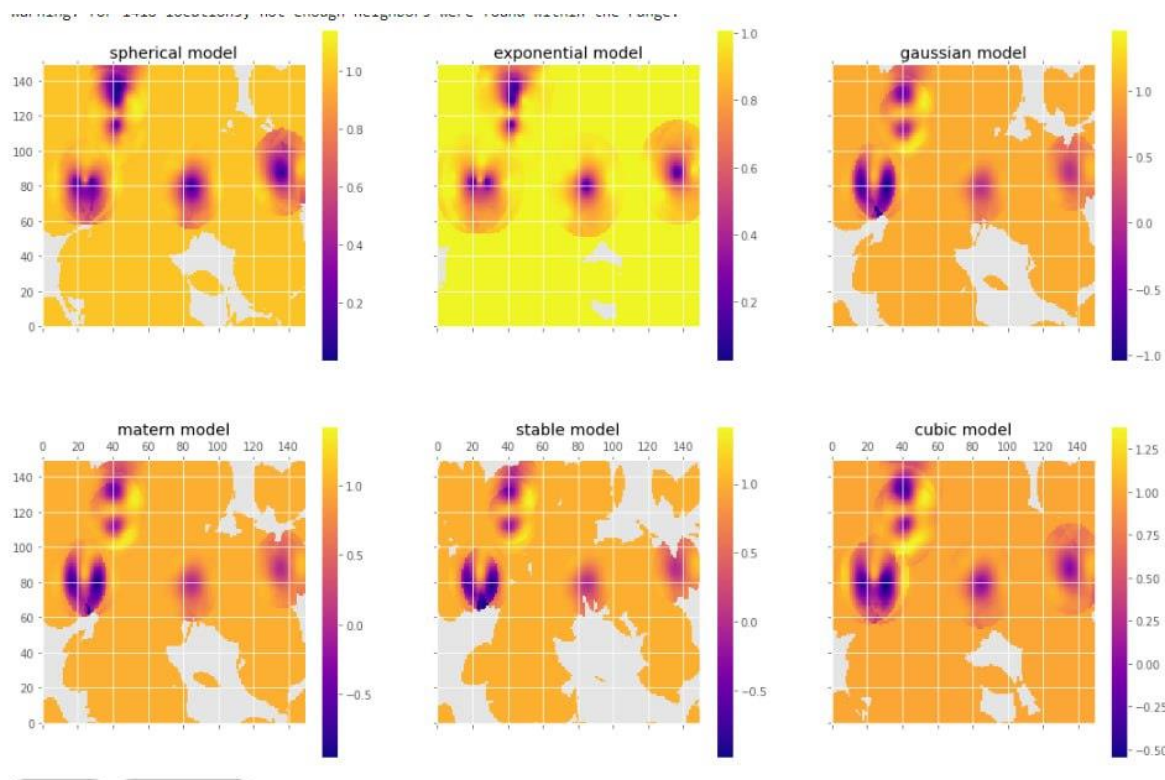


Рисунок 3.15 – Результат геостатистичного аналізу основними популярними моделями

Загалом роботу інформаційної технології та відповідальність кожного з присутніх модулів можна описати UML діаграмою компонентів що зображена у Додатку В.

3.4 Висновки

На основі зібраних даних спостережень робимо висновок що внаслідок геостатистичного аналізу найточніші показники можна помітити якраз у країнах, де станцій найбільше та розташовані вони найщільніше - це Німеччина, Австрія, Бельгія та Нідерланди. Найкраще ситуацію відображають «matern» та «stable» моделі.

З графіків можна чітко прослідкувати рівень заміру трьох основних нуклеотидів та рівні його розповсюдження що здійснювались протягом року після аварії на ЧАЕС. Також вдалось помітити певну закономірність - практично всюди рівень забруднення розсіявся рівномірно за виключенням міст та агломерацій що є сильно озелененими та мають дещо нижчий рівень урбанізації (для прикладу, Берлін, Гамбург, Париж, Амстердам тощо) - це й помітно на графіках специфічними потемніннями. «Чисті» зони можна з легкістю помітити й на мапах наданих джерелами Об'єднаного дослідного центру ЄС що описувались попередньо у першому розділі. Тому можемо сміливо робити висновок, що геостатистичний аналіз виконаний з винятковою точністю.

На основі проведених досліджень та аналізу зібраних даних можна виділити кілька важливих пунктів:

1. Ефективна підготовка даних є критичною для точності результатів. Нормалізація, очищення та заповнення відсутніх значень дозволяє забезпечити коректну інтерпретацію даних та уникнути помилок у подальшому аналізі.
2. Методи кластеризації та інтерполяції допомагають виявити закономірності у поширенні радіаційного забруднення. Завдяки цьому можна більш точно інтерполювати рівень забруднення та ефективніше планувати заходи з його зниження.
3. Геостатистичний аналіз за допомогою варіограм і методу кригінг дозволяє проводити просторову інтерполяцію та передбачати рівень радіації у місцях, де вимірювання не проводилися. Це значно підвищує ефективність моніторингу та дозволяє покращити якість аналітичних спостережень радіаційного фону на великих територіях.
4. Візуалізація даних за допомогою карт, графіків та варіограм є важливим інструментом для прийняття рішень. Вона дозволяє представити результати досліджень у наочній формі, що полегшує аналіз та інтерпретацію даних.

Загалом, поєднання сучасних методів аналізу даних, таких як інтерполяція,

кластеризація та геостатистичний аналіз, дозволяє підвищити точність прогнозування радіаційного фону та зробити моніторинг радіаційного забруднення країн Європи достатньо ефективним для того, щоб використовувати такий підхід як в безпекових, так і в комерційних цілях.

4 ЕКОНОМІЧНА ЧАСТИНА

Виконання науково-дослідної роботи завжди передбачає отримання певних результатів і вимагає відповідних витрат. Результати виконаної роботи завжди дають нам нові знання, які в подальшому можуть бути використані для удосконалення та/або розробки (побудови) нових, більш продуктивних зразків техніки, процесів та програмного забезпечення.

Дослідження на тему «Інформаційна технологія геостатистичного аналізу радіаційного забруднення в Європі» може бути віднесено до фундаментальних і пошукових наукових досліджень і спрямоване на вирішення наукових проблем, пов'язаних з практичним застосуванням. Основою таких досліджень є науковий ефект, який виражається в отриманні наукових результатів, які збільшують обсяг знань про природу, техніку та суспільство, які розвивають теоретичну базу в тому чи іншому науковому напрямку, що дозволяє виявити нові закономірності, які можуть використовуватися на практиці.

Для цього випадку виконаємо такі етапи робіт:

- 1) здійснимо проведення наукового аудиту досліджень, тобто встановлення їх наукового рівня та значимості;
- 2) проведемо планування витрат на проведення наукових досліджень;
- 3) здійснимо розрахунок рівня важливості наукового дослідження та перспективності, визначимо ефективність наукових досліджень.

4.1 Оцінювання наукового ефекту

Основними ознаками наукового ефекту науково-дослідної роботи є новизна роботи, рівень її теоретичного опрацювання, перспективність, рівень розповсюдження результатів, можливість реалізації. Науковий ефект НДР на тему «Інформаційна технологія геостатистичного аналізу радіаційного забруднення в Європі» можна охарактеризувати двома показниками: ступенем наукової новизни та рівнем теоретичного опрацювання.

Значення показників ступеня новизни і рівня теоретичного опрацювання науково-дослідної роботи в балах наведені в таблицях 4.1 та 4.2.

Таблиця 4.1 – Показники ступеня новизни науково-дослідної роботи виставлені експертами

Ступінь новизни	Характеристика ступеня новизни	Значення ступеня новизни, бали		
		Експерти (ПІБ, посада)		
		1	2	3
Принципово нова	Робота якісно нова за постановкою задачі і ґрунтується на застосуванні оригінальних методів дослідження. Результати дослідження відкривають новий напрям в даній галузі науки і техніки. Отримані принципово нові факти, закономірності; розроблена нова теорія. Створено принципово новий пристрій, спосіб, метод	0	0	0
Нова	Отримана нова інформація, яка суттєво зменшує невизначеність наявних значень (по-новому або вперше пояснені відомі факти, закономірності, впроваджені нові поняття, розкрита структура змісту). Проведено суттєве вдосконалення, доповнення і уточнення раніше досягнутих результатів	50	50	51
Відносно нова	Робота має елементи новизни в постановці задачі і методах дослідження. Результати дослідження систематизують і узагальнюють наявну інформацію, визначають шляхи подальших досліджень; вперше знайдено зв'язок (або знайдено новий зв'язок) між явищами. В принципі відомі положення розповсюджені на велику кількість об'єктів, в результаті чого знайдено ефективне рішення. Розроблені більш прості способи для досягнення відомих результатів. Проведена часткова раціональна модифікація (з ознаками новизни)	0	0	0
Традиційна	Робота виконана за традиційною методикою. Результати дослідження мають інформаційний характер. Підтверджені або поставлені під сумнів відомі факти та твердження, які потребують перевірки. Знайдено новий варіант рішення, який не дає суттєвих переваг в порівнянні з існуючим	0	0	0
Не нова	Отримано результат, який раніше зафіксований в інформаційному полі, та не був відомий авторам	0	0	0
Середнє значення балів експертів		50,3		

Згідно отриманого середнього значення балів експертів ступінь новизни характеризується як нова, тобто отримана нова інформація, яка суттєво зменшує невизначеність наявних знань (по-новому або вперше пояснені відомі факти, закономірності, впроваджені нові поняття, розкрита структура змісту) та проведено суттєве вдосконалення, доповнення і уточнення раніше досягнутих результатів.

Таблиця 4.2 – Показники рівня теоретичного опрацювання науково-дослідної роботи виставлені експертами

Характеристика рівня теоретичного опрацювання	Значення показника рівня теоретичного опрацювання, бали		
	Експерт (ПІБ, посада)		
	1	2	3
Відкриття закону, розробка теорії	0	0	0
Глибоке опрацювання проблеми: багатоаспектний аналіз зв'язків, взаємозалежності між фактами з наявністю пояснень, наукової систематизації з побудовою евристичної моделі або комплексного прогнозу	60	62	60
Розробка способу (алгоритму, програми), пристрою, отримання нової речовини	0	0	0
Елементарний аналіз зв'язків між фактами та наявною гіпотезою, класифікація, практичні рекомендації для окремого випадку тощо	0	0	0
Опис окремих елементарних фактів, викладення досвіду, результатів спостережень, вимірювань тощо	0	0	0
Середнє значення балів експертів	60,7		

Згідно отриманого середнього значення балів експертів рівень теоретичного опрацювання науково-дослідної роботи характеризується як глибоке опрацювання проблеми: багатоаспектний аналіз зв'язків, взаємозалежності між фактами з наявністю пояснень, наукової систематизації з побудовою евристичної моделі або комплексного прогнозу.

Показник, який характеризує рівень наукового ефекту, визначаємо за формулою [29]:

$$E_{\text{нау}} = 0,6 \cdot k_{\text{нов}} + 0,4 \cdot k_{\text{теор}}, \quad (4.1)$$

де $k_{\text{нов}}, k_{\text{теор}}$ - показники ступеня новизни та рівня теоретичного опрацювання науково-дослідної роботи,

$$k_{\text{нов}} = 50,3,$$

$$k_{\text{теор}} = 60,7 \text{ балів};$$

0,6 та 0,4 – питома вага (значимість) показників ступеня новизни та рівня теоретичного опрацювання науково-дослідної роботи.

$$E_{\text{нау}} = 0,6 \cdot k_{\text{нов}} + 0,4 \cdot k_{\text{теор}} = 0,6 \cdot 50,3 + 0,4 \cdot 60,67 = 54,47 \text{ балів.}$$

Визначення характеристики показника $E_{\text{нау}}$ проводиться на основі висновків експертів виходячи з граничних значень, які наведені в таблиці 4.3.

Таблиця 4.3 – Граничні значення показника наукового ефекту

Досягнутий рівень показника	Кількість балів
Високий	70...100
Середній	50...69
Достатній	15...49
Низький (помилкові дослідження)	1...14

Відповідно до визначеного рівня наукового ефекту проведеної науково-дослідної роботи на тему «Інформаційна технологія геостатистичного аналізу радіаційного забруднення в Європі», даний рівень становить 54,47 балів і відповідає статусу - середній рівень. Тобто у даному випадку можна вести мову про потенційну фактичну ефективність науково-дослідної роботи.

4.2 Розрахунок витрат на здійснення науково-дослідної роботи

Витрати, пов'язані з проведенням науково-дослідної роботи на тему «Інформаційна технологія геостатистичного аналізу радіаційного забруднення в Європі», під час планування, обліку і калькулювання собівартості науково-дослідної роботи групуємо за відповідними статтями.

4.2.1 Витрати на оплату праці

До статті «Витрати на оплат у праці» належать витрати на виплату основної та додаткової заробітної плати керівникам відділів, лабораторій, секторів і груп, науковим, інженерно-технічним працівникам, конструкторам, технологам, креслярам, копіювальникам, лаборантам, робітникам, студентам, аспірантам та іншим працівникам, безпосередньо зайнятим виконанням конкретної теми, обчисленої за посадовими окладами, відрядними розцінками, тарифними ставками згідно з чинними в організаціях системами оплати праці.

Основна заробітна плата дослідників

Витрати на основну заробітну плату дослідників ($З_o$) розраховуємо у відповідності до посадових окладів працівників, за формулою [29]:

$$З_o = \sum_{i=1}^k \frac{M_{ni} \cdot t_i}{T_p}, \quad (4.2)$$

де k – кількість посад дослідників залучених до процесу досліджень;

M_{ni} – місячний посадовий оклад конкретного дослідника, грн;

t_i – число днів роботи конкретного дослідника, дн.;

T_p – середнє число робочих днів в місяці, $T_p=21$ дні.

$$З_o = 18550,00 \cdot 18 / 21 = 15900,00 \text{ грн.}$$

Проведені розрахунки зведемо до таблиці 4.4.

Таблиця 4.4 – Витрати на заробітну плату дослідників

Найменування посади	Місячний посадовий оклад, грн	Оплата за робочий день, грн	Число днів роботи	Витрати на заробітну плату, грн
Керівник науково-дослідної роботи	18550,00	883,33	18	15900,00
Інженер-розробник програмного забезпечення формування веб-сайту	17850,00	850,00	36	30600,00
Консультант (фахівець-еколог)	17000,00	809,52	5	4047,62
Технік 1-ї категорії	8150,00	388,10	10	3880,95
Всього				54428,57

Основна заробітна плата робітників

Витрати на основну заробітну плату робітників ($З_p$) за відповідними найменуваннями робіт НДР на тему «Інформаційна технологія геостатистичного аналізу радіаційного забруднення в Європі» розраховуємо за формулою:

$$З_p = \sum_{i=1}^n C_i \cdot t_i, \quad (4.3)$$

де C_i – погодинна тарифна ставка робітника відповідного розряду, за виконану відповідну роботу, грн/год;

t_i – час роботи робітника при виконанні визначеної роботи, год.

Погодинну тарифну ставку робітника відповідного розряду C_i можна визначити за формулою:

$$C_i = \frac{M_M \cdot K_i \cdot K_c}{T_p \cdot t_{зм}}, \quad (4.4)$$

де M_M – розмір прожиткового мінімуму працездатної особи, або мінімальної місячної заробітної плати (в залежності від діючого законодавства), прийmemo $M_M=8000,00$ грн;

K_i – коефіцієнт міжкваліфікаційного співвідношення для встановлення тарифної ставки робітнику відповідного розряду [29];

K_c – мінімальний коефіцієнт співвідношень місячних тарифних ставок робітників першого розряду з нормальними умовами праці виробничих об'єднань і підприємств до законодавчо встановленого розміру мінімальної заробітної плати.

T_p – середнє число робочих днів в місяці, приблизно $T_p = 21$ дн;

$t_{зм}$ – тривалість зміни, год.

$$C_l = 8000,00 \cdot 1,10 \cdot 1,15 / (21 \cdot 8) = 60,24 \text{ грн.}$$

$$З_{pl} = 60,24 \cdot 3,50 = 210,83 \text{ грн.}$$

Таблиця 4.5 – Величина витрат на основну заробітну плату робітників

Найменування робіт	Тривалість роботи, год	Розряд роботи	Тарифний коефіцієнт	Погодинна тарифна ставка, грн	Величина оплати на робітника грн
Розміщення комп'ютерного обладнання розробки ПЗ	3,50	2	1,10	60,24	210,83
Інсталяція програмного забезпечення розробки	5,50	4	1,50	82,14	451,79
Підготовка робочого місця розробника програмних засобів	4,56	3	1,35	73,93	337,11
Формування бази даних радіаційного забруднення	15,00	2	1,10	60,24	903,57
Компіляція програмних блоків	6,00	5	1,70	93,10	558,57
Тестування ПЗ	6,00	2	1,10	60,24	361,43
Всього					2823,30

Додаткова заробітна плата дослідників та робітників

Додаткову заробітну плату розраховуємо як 10 ... 12% від суми основної заробітної плати дослідників та робітників за формулою:

$$Z_{\text{доо}} = (Z_o + Z_p) \cdot \frac{H_{\text{доо}}}{100\%}, \quad (4.5)$$

де $H_{\text{доо}}$ – норма нарахування додаткової заробітної плати. Прийmemo 11%.

$$Z_{\text{доо}} = (54428,57 + 2823,30) \cdot 11 / 100\% = 6297,71 \text{ грн.}$$

4.2.2 Відрахування на соціальні заходи

Нарахування на заробітну плату дослідників та робітників розраховуємо як 22% від суми основної та додаткової заробітної плати дослідників і робітників за формулою:

$$Z_n = (Z_o + Z_p + Z_{\text{доо}}) \cdot \frac{H_{\text{зн}}}{100\%} \quad (4.6)$$

де $H_{\text{зн}}$ – норма нарахування на заробітну плату. Приймаємо 22%.

$$Z_n = (54428,57 + 2823,30 + 6297,71) \cdot 22 / 100\% = 13980,91 \text{ грн.}$$

4.2.2 Сировина та матеріали

До статті «Сировина та матеріали» належать витрати на сировину, основні та допоміжні матеріали, інструменти, пристрої та інші засоби і предмети праці, які придбані у сторонніх підприємств, установ і організацій та витрачені на проведення досліджень за темою «Інформаційна технологія геостатистичного аналізу радіаційного забруднення в Європі».

Витрати на матеріали на даному етапі проведення досліджень в основному пов'язані з використанням моделей елементів та моделювання роботи і досліджень за допомогою комп'ютерної техніки та створення експериментальних математичних моделей або програмного забезпечення, тому дані витрати формуються на основі витратних матеріалів характерних для офісних робіт.

Витрати на матеріали (M), у вартісному вираженні розраховуються окремо

по кожному виду матеріалів за формулою:

$$M = \sum_{j=1}^n H_j \cdot C_j \cdot K_j - \sum_{j=1}^n B_j \cdot C_{ej}, \quad (4.7)$$

де H_j – норма витрат матеріалу j -го найменування, кг;

n – кількість видів матеріалів;

C_j – вартість матеріалу j -го найменування, грн/кг;

K_j – коефіцієнт транспортних витрат, ($K_j = 1,1 \dots 1,15$);

B_j – маса відходів j -го найменування, кг;

C_{ej} – вартість відходів j -го найменування, грн/кг.

$$M_1 = 3,0 \cdot 225,00 \cdot 1,1 - 0 \cdot 0 = 742,50 \text{ грн.}$$

Проведені розрахунки зведемо до таблиці 4.6.

Таблиця 4.6 – Витрати на матеріали

Найменування матеріалу, марка, тип, сорт	Ціна за 1 кг, грн	Норма витрат, кг	Величина відходів, кг	Ціна відходів, грн/кг	Вартість витраченого матеріалу, грн
Папір канцелярський офісний (A4)	225,00	3,0	0	0	742,50
Папір для заміток (A5)	116,00	3,0	0	0	382,80
Начиння канцелярське	195,00	3,0	0	0	643,50
Органайзер офісний	183,00	3,0	0	0	603,90
Картридж для принтера	950,00	2,0	0	0	2090,00
Диск оптичний	25,00	6,0	0	0	165,00
USB-пам'ять Microtech 64 GB	199,00	2,0	0	0	437,80
Тека для паперів BBS-QT	89,90	3,0	0	0	296,67
Всього					5362,17

4.2.3 Розрахунок витрат на комплектуючі

Витрати на комплектуючі (K_6), які використовують при проведенні НДР на тему «Інформаційна технологія геостатистичного аналізу радіаційного забруднення в Європі», розраховуємо, згідно з їхньою номенклатурою, за формулою:

$$K_6 = \sum_{j=1}^n H_j \cdot C_j \cdot K_j \quad (4.8)$$

де H_j – кількість комплектуючих j -го виду, шт.;

C_j – покупна ціна комплектуючих j -го виду, грн;

K_j – коефіцієнт транспортних витрат, ($K_j = 1,1 \dots 1,15$).

$K_6 = 1 \cdot 1599,00 \cdot 1,1 = 1758,90$ грн.

Проведені розрахунки зведемо до таблиці 4.7.

Таблиця 4.7 – Витрати на комплектуючі

Найменування комплектуючих	Кількість, шт.	Ціна за штуку, грн	Сума, грн
Зовнішній жорсткий диск 2.5" 1TB Seagate (STGD2000200)	1	1599,00	1758,90
Концентратор Defender SEPTIMA SLIM (83505)	1	850,00	935,00
Кабель для передачі даних USB to COM 1.0m Patron (CAB-PN-USB-COM)	1	420,00	462,00
Всього			3155,90

4.2.5 Спецустаткування для наукових (експериментальних) робіт

До статті «Спецустаткування для наукових (експериментальних) робіт» належать витрати на виготовлення та придбання спецустаткування необхідного для проведення досліджень, також витрати на їх проектування, виготовлення, транспортування, монтаж та встановлення.

Балансову вартість спецустаткування розраховуємо за формулою:

$$B_{\text{спец}} = \sum_{i=1}^k C_i \cdot C_{\text{пр.і}} \cdot K_i, \quad (4.9)$$

де C_i – ціна придбання одиниці спецустаткування даного виду, марки, грн;
 $C_{\text{пр.і}}$ – кількість одиниць устаткування відповідного найменування, які придбані для проведення досліджень, шт.;

K_i – коефіцієнт, що враховує доставку, монтаж, налагодження устаткування тощо, ($K_i = 1,10 \dots 1,12$);

k – кількість найменувань устаткування.

$$B_{\text{спец}} = 2499,00 \cdot 1 \cdot 1,1 = 2748,90 \text{ грн.}$$

Отримані результати зведемо до таблиці 4.8:

Таблиця 4.8 – Витрати на придбання спецустаткування по кожному виду

Найменування устаткування	Кількість, шт	Ціна за одиницю, грн	Вартість, грн
Wi-Fi роутер Xiaomi Router AX3200 (DVB4314GL)	1	2499,00	2748,90
Всього			2748,90

4.2.6 Програмне забезпечення для наукових робіт

До статті «Програмне забезпечення для наукових (експериментальних) робіт» належать витрати на розробку та придбання спеціальних програмних засобів і програмного забезпечення, (програм, алгоритмів, баз даних) необхідних для проведення досліджень, також витрати на їх проектування, формування та встановлення.

Балансову вартість програмного забезпечення розраховуємо за формулою:

$$B_{npz} = \sum_{i=1}^k C_{inpr} \cdot C_{npz.i} \cdot K_i, \quad (4.10)$$

де C_{inpr} – ціна придбання одиниці програмного засобу даного виду, грн;
 $C_{npz.i}$ – кількість одиниць програмного забезпечення відповідного найменування, які придбані для проведення досліджень, шт.;

K_i – коефіцієнт, що враховує інсталяцію, налагодження програмного засобу тощо, ($K_i = 1, 10 \dots 1, 12$);

k – кількість найменувань програмних засобів.

$$B_{npz} = 40,38 \cdot 1 \cdot 1,1 = 44,42 \text{ грн.}$$

Отримані результати зведемо до таблиці 4.9.

Таблиця 4.9 – Витрати на придбання програмних засобів по кожному виду

Найменування програмного засобу	Кількість, шт	Ціна за одиницю, грн	Вартість, грн
Середовище розробки Visual Studio	1	40,38	44,42
Хмарний хостинг Cloud Professional для отримання даних (доступ, місяць)	2	480,00	1056,00
Абонентна плата доступу до мережі Internet (за місяць)	2	382,00	840,40
ArcGIS Online Developer API token	1	40,38	44,42
Всього			1985.23

4.2.7 Амортизація обладнання, програмних засобів та приміщень

В спрощеному вигляді амортизаційні відрахування по кожному виду обладнання, приміщень та програмному забезпеченню тощо, розраховуємо з використанням прямолінійного методу амортизації за формулою:

$$A_{обл} = \frac{Ц_{б}}{T_{в}} \cdot \frac{t_{вик}}{12}, \quad (4.11)$$

де $Ц_{б}$ – балансова вартість обладнання, програмних засобів, приміщень тощо, які використовувались для проведення досліджень, грн;

$t_{вик}$ – термін використання обладнання, програмних засобів, приміщень під час досліджень, місяців;

$T_{в}$ – строк корисного використання обладнання, програмних засобів, приміщень тощо, років.

$$A_{обл} = (42899,00 \cdot 2) / (2 \cdot 12) = 3574,92 \text{ грн.}$$

Проведені розрахунки зведемо до таблиці 4.10.

Таблиця 4.10 – Амортизаційні відрахування по кожному виду обладнання

Найменування обладнання	Балансова вартість, грн	Строк корисного використання, років	Термін використання обладнання, місяців	Амортизаційні відрахування, грн
Персональний комп'ютер (AMD Ryzen 6800HS, Nvidia GeForce RTX 3050, 16GB)	42899,00	2	2	3574,92
Робоче місце розробника ПЗ	8570,00	5	2	285,67
Пристрої виводу	9799,00	4	2	408,29
Оргтехніка	9240,00	4	2	385,00
Приміщення лабораторії розробки та дослідження інформаційної технології	370000,00	25	2	2466,67
ОС Windows 11	5800,00	2	2	483,33
Прикладний пакет Microsoft Office 2021	5100,00	2	2	425,00
Обладнання передачі цифрових даних	9650,00	5	2	321,67
Всього				8350,54

4.2.8 Паливо та енергія для науково-виробничих цілей

Витрати на силову електроенергію (B_e) розраховуємо за формулою:

$$B_e = \sum_{i=1}^n \frac{W_{yi} \cdot t_i \cdot C_e \cdot K_{\text{вн}i}}{\eta_i}, \quad (4.12)$$

де W_{yi} – встановлена потужність обладнання на визначеному етапі розробки, кВт;

t_i – тривалість роботи обладнання на етапі дослідження, год;

C_e – вартість 1 кВт-години електроенергії, грн; (вартість електроенергії визначається за даними енергопостачальної компанії), прийmemo $C_e = 10,98$ грн;

$K_{\text{вн}i}$ – коефіцієнт, що враховує використання потужності, $K_{\text{вн}i} < 1$;

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

$$B_e = 0,42 \cdot 160,0 \cdot 10,98 \cdot 0,95 / 0,97 = 737,86 \text{ грн.}$$

Проведені розрахунки зведемо до таблиці.

Таблиця 4.11 – Витрати на електроенергію

Найменування обладнання	Встановлена потужність, кВт	Тривалість роботи, год	Сума, грн
Персональний комп'ютер Core I5-9600	0,42	160,0	737,86
Робоче місце розробника ПЗ	0,12	240,0	316,22
Пристрої виводу інформації	0,20	12,0	26,35
Оргтехніка	0,40	4,0	17,57
Сервер HP Proliant DL 360 Gen10 (8x2.5) SFF	0,25	160,0	439,20
Wi-Fi роутер Xiaomi Router AX3200	0,03	240,0	79,06
Всього			1616,26

4.2.9 Службові відрядження

До статті «Службові відрядження» дослідної роботи на тему «Інформаційна технологія геостатистичного аналізу радіаційного забруднення в Європі» належать витрати на відрядження штатних працівників, працівників організацій, які працюють за договорами цивільно-правового характеру, аспірантів, зайнятих розробленням досліджень, відрядження, пов'язані з проведенням випробувань машин та приладів, а також витрати на відрядження на наукові з'їзди, конференції, наради, пов'язані з виконанням конкретних досліджень.

Витрати за статтею «Службові відрядження» розраховуємо як 20...25% від суми основної заробітної плати дослідників та робітників за формулою:

$$B_{cv} = (Z_o + Z_p) \cdot \frac{H_{cv}}{100\%}, \quad (4.13)$$

де H_{cv} – норма нарахування за статтею «Службові відрядження», приймемо $H_{cv} = 20\%$.

$$B_{cv} = (54428,57 + 2823,30) \cdot 20 / 100\% = 11450,38 \text{ грн.}$$

4.2.10 Витрати на роботи, які виконують сторонні підприємства, установи і організації

Витрати за статтею «Витрати на роботи, які виконують сторонні підприємства, установи і організації» розраховуємо як 30...45% від суми основної заробітної плати дослідників та робітників за формулою:

$$B_{cn} = (Z_o + Z_p) \cdot \frac{H_{cn}}{100\%}, \quad (4.14)$$

де H_{cn} – норма нарахування за статтею «Витрати на роботи, які виконують сторонні підприємства, установи і організації», прийmemo $H_{cn} = 30\%$.

$$B_{cn} = (54428,57 + 2823,30) \cdot 30 / 100\% = 17175,56 \text{ грн.}$$

4.2.11 Інші витрати

До статті «Інші витрати» належать витрати, які не знайшли відображення у зазначених статтях витрат і можуть бути віднесені безпосередньо на собівартість досліджень за прямими ознаками.

Витрати за статтею «Інші витрати» розраховуємо як 50...100% від суми основної заробітної плати дослідників та робітників за формулою:

$$I_{is} = (Z_o + Z_p) \cdot \frac{H_{is}}{100\%}, \quad (4.15)$$

де H_{is} – норма нарахування за статтею «Інші витрати», прийmemo $H_{is} = 50\%$.

$$I_{is} = (54428,57 + 2823,30) \cdot 50 / 100\% = 28625,94 \text{ грн.}$$

4.2.12 Накладні (загальновиробничі) витрати

До статті «Накладні (загальновиробничі) витрати» належать: витрати, пов'язані з управлінням організацією; витрати на винахідництво та раціоналізацію; витрати на підготовку (перепідготовку) та навчання кадрів; витрати, пов'язані з набором робочої сили; витрати на оплату послуг банків; витрати, пов'язані з освоєнням виробництва продукції; витрати на науково-технічну інформацію та рекламу та ін.

Витрати за статтею «Накладні (загальновиробничі) витрати» розраховуємо як 100...150% від суми основної заробітної плати дослідників та робітників за формулою:

$$B_{nzb} = (Z_o + Z_p) \cdot \frac{H_{nzb}}{100\%}, \quad (4.16)$$

де $H_{нзв}$ – норма нарахування за статтею «Накладні (загальноновиробничі) витрати», прийmemo $H_{нзв} = 100\%$.

$$B_{нзв} = (54428,57 + 2823,30) \cdot 100 / 100\% = 57251,88 \text{ грн.}$$

Витрати на проведення науково-дослідної роботи на тему «Інформаційна технологія геостатистичного аналізу радіаційного забруднення в Європі» розраховуємо як суму всіх попередніх статей витрат за формулою:

$$B_{заг} = Z_o + Z_p + Z_{доо} + Z_n + M + K_v + B_{спец} + B_{прз} + A_{обл} + B_e + B_{св} + B_{сп} + I_v + B_{нзв}. \quad (4.17)$$

$$B_{заг} = 54428,57 + 2823,30 + 6297,71 + 13980,91 + 5362,17 + 3155,90 + 2748,90 + 1985,23 + 8350,54 + 1616,26 + 11450,38 + 17175,56 + 28625,94 + 57251,88 = 215253.24 \text{ грн.}$$

Загальні витрати $ЗВ$ на завершення науково-дослідної (науково-технічної) роботи та оформлення її результатів розраховується за формулою:

$$ЗВ = \frac{B_{заг}}{\eta}, \quad (4.18)$$

де η - коефіцієнт, який характеризує етап (стадію) виконання науково-дослідної роботи, прийmemo $\eta = 0,9$.

$$ЗВ = 215253.24 / 0,9 = 239170.267 \text{ грн.}$$

4.3 Оцінювання важливості та наукової значимості науково-дослідної роботи

Оцінювання та доведення ефективності виконання науково-дослідної роботи фундаментального чи пошукового характеру є достатньо складним процесом і часто базується на експертних оцінках, тому має вірогідний характер.

Для обґрунтування доцільності виконання науково-дослідної роботи на тему «Інформаційна технологія геостатистичного аналізу радіаційного

забруднення в Європі» використовується спеціальний комплексний показник, що враховує важливість, результативність роботи, можливість впровадження її результатів у виробництво, величину витрат на роботу.

Комплексний показник K_p рівня науково-дослідної роботи може бути розрахований за формулою:

$$K_p = \frac{I^n \cdot T_C \cdot R}{B \cdot t}, \quad (4.19)$$

де I – коефіцієнт важливості роботи. Прийmemo $I = 4$;

n – коефіцієнт використання результатів роботи; $n = 0$, коли результати роботи не будуть використовуватись; $n = 1$, коли результати роботи будуть використовуватись частково; $n = 2$, коли результати роботи будуть використовуватись в дослідно-конструкторських розробках; $n = 3$, коли результати можуть використовуватись навіть без проведення дослідно-конструкторських розробок. Прийmemo $n = 3$;

T_C – коефіцієнт складності роботи. Прийmemo $T_C = 2$;

R – коефіцієнт результативності роботи; якщо результати роботи плануються вище відомих, то $R = 4$; якщо результати роботи відповідають відомому рівню, то $R = 3$; якщо нижче відомих результатів, то $R = 1$. Прийmemo $R = 4$;

B – вартість науково-дослідної роботи, тис. грн. Прийmemo $B = 239268,98$ грн;

t – час проведення дослідження. Прийmemo $t = 0,17$ років, (2 міс.).

Визначення показників I, n, T_C, R, B, t здійснюється експертним шляхом або на основі нормативів [29].

$$K_p = \frac{I^n \cdot T_C \cdot R}{B \cdot t} = 4^3 \cdot 2 \cdot 4 / 239,2 \cdot 0,17 = 12,59.$$

Якщо $K_p > 1$, то науково-дослідну роботу на тему «Інформаційна технологія

геостатистичного аналізу радіаційного забруднення в Європі» можна вважати ефективною з високим науковим, технічним і економічним рівнем.

4.4 Висновок

Витрати на проведення науково-дослідної роботи на тему «Інформаційна технологія геостатистичного аналізу радіаційного забруднення в Європі» складають 239170.267 грн. Відповідно до проведеного аналізу та розрахунків рівень наукового ефекту проведеної науково-дослідної роботи на тему «Інформаційна технологія геостатистичного аналізу радіаційного забруднення в Європі» є середній, а дослідження актуальними, рівень доцільності виконання науково-дослідної роботи $K_p > 1$, що свідчить про потенційну ефективність з високим науковим, технічним і економічним рівнем.

ВИСНОВКИ

У ході виконання даної кваліфікаційної роботи було проведено дослідження щодо радіаційного забруднення в країнах Європи після аварії на Чорнобильській АЕС. Використання сучасних інформаційних технологій, таких як геоінформаційні системи (ГІС), методи геостатистичного аналізу та застосування методів машинного навчання, дозволило ефективно обробляти та аналізувати великі обсяги даних про радіаційне забруднення.

Основними методами, які використовувалися у дослідженні, були інтерполяція даних, використання геостатистичних моделей та аналіз просторового розподілу радіаційного забруднення. Застосування цих методів дозволило побудувати карту забруднення, яка відображає рівень радіаційного забруднення на території Європи після аварії на ЧАЕС.

Проведене дослідження показало, що найкращі результати для даної задачі було досягнуто за рахунок використання геостатистичного методу кригінгу для аналізу просторових даних на основі моделей «matern» та «stable», що дозволяє покращити точність та ефективність моніторингу, прогнозування та управління радіаційними ризиками в Європі

У роботі були досягнуті всі поставлені завдання, що дозволило виконати комплексний аналіз радіаційного забруднення в Європі. Зокрема:

- Проведено аналіз сучасних інформаційних технологій для обробки просторових даних, що дало змогу вибрати оптимальні середовища та інструменти для розробки.
- Визначено та інтегровано джерела даних про радіаційне забруднення, включаючи рівні радіаційного фону та концентрації радіонуклідів, з використанням методів лінійної інтерполяції для їх первинної обробки.
- Виконано оброблення та підготовку даних для геостатистичного аналізу, що забезпечило їхню якість і готовність до використання.
- Застосовано інтерполяційні та геостатистичні методи, на основі яких створено карти забруднення.

- Проведено інтерпретацію отриманих результатів та їхню візуалізацію у вигляді картографічних матеріалів, графіків і наочних схем, що підвищило доступність даних для аналізу.

Результати виконаних завдань стали основою для реалізації ефективної інформаційної технології, спрямованої на вдосконалення моніторингу та аналізу радіаційного забруднення, що відповідає поставленій меті роботи.

Використання інструментів, таких як ArcGIS та Python з відповідними бібліотеками, сприяло удосконаленню процесу моніторингу якості радіаційного забруднення. Отримані результати свідчать про важливість використання сучасних підходів та технологій у аналізі та інтерполяції даних радіаційного забруднення для забезпечення ефективного контролю над екологічною ситуацією в Європі.

Проведене дослідження підкреслює важливість інтеграції сучасних інформаційних технологій у процеси моніторингу радіаційного забруднення. Геоінформаційні системи (ГІС), Big Data та алгоритми машинного навчання значно підвищують точність і швидкість обробки даних, що є ключовими факторами для забезпечення оперативного реагування на потенційні радіаційні загрози. Сучасні технології дозволяють не лише збирати великі обсяги даних у режимі реального часу, а й ефективно аналізувати їх для прогнозування поширення радіоактивних речовин та планування заходів з мінімізації ризиків для населення та довкілля.

У цілому, ця робота несе за собою значний практичний внесок у галузі екологічного моніторингу та контролю радіаційного забруднення. Використання сучасних методів аналізу даних у поєднанні з інформаційними технологіями дозволяє ефективно вирішувати складні завдання, пов'язані з оцінкою та прогнозуванням впливу радіаційного забруднення, що є критичним для забезпечення сталого розвитку та збереження навколишнього середовища.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Opanasiuk Bohdan. Geostatistical analysis of nuclear pollution in EU with SciKit-Gstat, 2024. URL: <https://www.kaggle.com/code/auserox/geostatistical-analysis-of-nuclear-pollution-in-eu> (дата звернення: 22.09.2024).
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. Highfield, Linda, et al. "Interactive web-based mapping: bridging technology and data for health." *International Journal of Health Geographics* 10 (2011): 1-13a. URL: <https://ij-healthgeographics.biomedcentral.com/articles/10.1186/1476-072X-10-69> (дата звернення: 06.10.2024)
4. Khopkar, S. M. *Environmental pollution monitoring and control*. New Age International, 2007. URL: https://books.google.com.ua/books?id=TAk21grzDZgC&printsec=copyright&redir_esc=y#v=onepage&q&f=false (дата звернення – 05.10.2024)
5. Cressie N. The origins of kriging. *Math Geol.* 1990. Vol. 22. С. 239–252. URL: <https://doi.org/10.1007/BF00889887> (дата звернення: 01.10.2024).
6. Goovaerts P. Kriging Interpolation. The Geographic Information Science & Technology Body of Knowledge. 2019. URL: <https://gistbok.ucgis.org/bok-topics/kriging-interpolation> (дата звернення: 05.12.2024).
7. Joint Research Centre. URL: <https://remon.jrc.ec.europa.eu/About> (дата звернення: 17.09.2024).
8. Li, Jin, and Andrew D. Heap. "Spatial interpolation methods applied in the environmental sciences: A review." *Environmental Modelling & Software* 53 (2014): 173-189. URL: <https://www.sciencedirect.com/science/article/abs/pii/S1364815213003113> (дата звернення: 25.09.2024)
9. Neteler, M., & Mitasova, H. (2002). Open source GIS: a GRASS GIS

approach (Vol. 689). *Springer Science & Business Media*. URL: https://books.google.com.ua/books/about/Open_Source_GIS.html?id=G20ECAAAQBAJ&redir_esc=y (дата звернення: 25.09.2024)

10. Yue P., Jiang L. BigGIS: How big data can shape next-generation GIS. 2014. URL: https://www.researchgate.net/publication/282724755_BigGIS_How_big_data_can_shape_next-generation_GIS (дата звернення: 30.09.2024).

11. 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.

12. 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.

13. Wes McKinney. Python for Data Analysis. URL: http://www.ir.juit.ac.in:8080/jspui/bitstream/123456789/6083/1/Python%20for%20Data%20Analysis_%20Data%20Wrangling%20with%20Pandas%2C%20NumPy%2C%20and%20IPython.pdf (дата звернення – 15.10.2024)

14. Li, Jin, and Andrew D. Heap. "Spatial interpolation methods applied in the environmental sciences: A review." *Environmental Modelling & Software* 53 (2014): 173-189. URL: <https://www.sciencedirect.com/science/article/abs/pii/S1364815213003113> (дата звернення: 25.09.2024)

15. 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.

16. 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.

17. Мокін В. Б., Дзюняк Д. Ю., Бондалетов 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. URL: <http://praci.vntu.edu.ua/index.php/praci/article/view/456/454> (дата звернення: 28.10.2024).

18. 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.

19. 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.

20. 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.

21. Опанасюк Б. М., Мокін В.Б., Розробка інформаційної технології геостатистичного аналізу радіаційного аналізу радіаційного забруднення. Всеукраїнська науково-практична інтернет-конференція «Молодь в науці: дослідження, проблеми, перспективи» (Вінниця, 2024-2025 рр.). URL: <https://conferences.vntu.edu.ua/index.php/>. (Дата звернення: 10.11.2024)

22. Мокін В. Б., Крижановський Є. М., Лебухорський А. І. Системний

ГІС-аналіз впливу асфальтування доріг та ремонту центрального моста на забруднення атмосферного повітря м. Вінниці у 2017 році. 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. URL: <https://conferences.vntu.edu.ua/index.php/all-fksa/all-fksa-2018/paper/view/4467/4420> (дата звернення: 04.10.2024).

23. Мокін В., Дзюняк Д. Технологія оцінювання параметрів викидів зі стаціонарних джерел за даними оперативного моніторингу. 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.

24. 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. Бондалстов, Г. В. Горячев, В. В. Боцяк : 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.

25. Giraudeau M., Bonzom J. M., Ducatez S., et al. Carotenoid distribution in wild Japanese tree frogs (*Hyla japonica*) exposed to ionizing radiation in Fukushima. *Sci Rep*. 2018. Vol. 8. С. 7438. URL: <https://doi.org/10.1038/s41598-018-25495-5> (дата звернення: 03.12.2024).

26. Copernicus project official website. URL: <https://www.copernicus.eu/en> (дата звернення: 10.10.2024).

27. ArcGIS Online Developer resources. URL: <https://www.esri.com/en-us/arcgis/products/arcgis-online/resources> (дата звернення: 28.09.2024).

28. Ying Sun, Bo Li, Marc G. Genton. *Geostatistics for Large Datasets*. 2012. URL: https://www.researchgate.net/publication/228807182_Geostatistics_for_Large_Data

[sets](#) (дата звернення: 07.11.2024).

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

Додаток А

Міністерство освіти і науки України
Вінницький національний технічний університет
Факультет інтелектуальних інформаційних технологій та автоматизації

ЗАТВЕРДЖУЮ

Завідувач кафедри САІТ

_____ д.т.н., проф. Віталій МОКІН

« ____ » _____ 2024 р.

ТЕХНІЧНЕ ЗАВДАННЯ

на магістерську кваліфікаційну роботу

«ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ГЕОСТАТИСТИЧНОГО АНАЛІЗУ
РАДІАЦІЙНОГО ЗАБРУДНЕННЯ В ЄВРОПІ»

08-34.МКР.002.02.000.ТЗ

Керівник: д.т.н. проф. САІТ

_____ Олександр МОКІН

« ____ » _____ 2024 р.

Розробив: студент гр. ІСТ-23м

_____ Богдан ОПАНАСЮК

« ____ » _____ 2024 р.

Вінниця 2024

1. Підстава для проведення робіт

Підставою для виконання роботи є наказ № __ по ВНТУ від «__» _____ 2024 р., та індивідуальне завдання на МКР, затверджене протоколом № __ засідання кафедри САІТ від «__» _____ 2024 р.

2. Джерела розробки:

– 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.

– 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.

– Мокін В. Б., Варчук І. В., Крижановський Є. М., Інформаційна технологія аналізу та оптимізації топологічної спостережуваності багатозв'язних геоінформаційних систем: монографія. 2019. URL: <https://press.vntu.edu.ua/index.php/vntu/catalog/book/522>.

3. Мета і призначення роботи:

Метою роботи є підвищення ефективності аналізу радіаційного забруднення в Європі шляхом створення інформаційної технології на основі геостатистичних методів.

4. Вихідні дані для проведення робіт:

- Kaggle Dataset «Chernobyl Chemical Radiation / CSV / Country Data» <https://www.kaggle.com/datasets/brsdincer/chernobyl-chemical-radiation-csv-country-data>

5. Методи дослідження:

- кластеризація даних
- інтерполяція даних
- кригінг-метод
- геостатистичний аналіз за допомогою SciKit-GStat

6. Етапи роботи і терміни їх виконання:

- а) Аналіз предметної області..... — —
- б) Вибір сучасних інструментів підходів до дослідження..... — —
- в) Розроблення інформаційної технології..... — —
- г) Економічна частина..... — —
- е) Оформлення пояснювальної записки — —

7. Очікувані результати та порядок реалізації:

Інформаційна технологія аналізу радіаційного забруднення в Європі, результатом якої є підвищення ефективності аналізу радіаційного забруднення в Європі шляхом створення інформаційної технології на основі геостатистичних методів.

8. Вимоги до розробленої документації

Текстова та ілюстративна частини роботи оформлені у відповідності до вимог «Методичних вказівок до виконання магістерських кваліфікаційних робіт для студентів спеціальності 126 «Інформаційні системи та технології» (освітня програма «Інформаційні технології аналізу даних та зображень»).

9. Порядок приймання роботи

Публічний захист «___» _____ 2024 р.

Початок розробки..... «___» _____ 2024 р.

Граничні терміни виконання МКР «___» _____ 2024 р.

Розробив студент групи ІСТ-23м _____ Богдан ОПАНАСЮК

Додаток Б

Протокол перевірки кваліфікаційної роботи на наявність текстових запозичень

Назва роботи: «Інформаційна технологія геостатистичного аналізу радіаційного забруднення в Європі»

Тип роботи: магістерська кваліфікаційна робота

Підрозділ: кафедра САІТ

Керівник: Олександр МОКІН, д.т.н., проф. каф. САІТ

Показники звіту подібності Turnitin

Оригінальність 94 %

Загальна схожість 6 %

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 відсутності 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 ознайомлений з повним звітом подібності, який був згенерований системою Turnitin щодо роботи.

Автор роботи

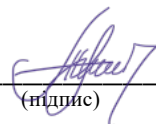

(підпис)

Богдан ОПАНАСЮК

Опис прийнятого рішення

Робота допускається до захисту

Особа, відповідальна за перевірку


(підпис)

Сергій ЖУКОВ

Додаток В

Лістинг програми

```

# Preprocessing: Handle missing values
# Option 1: Drop rows with any missing values in relevant columns
df_clean = df.dropna(subset=['Iodine_131', 'Caesium_134', 'Caesium_137'])

# Option 2: Fill missing values with the mean (or median) of the column
# df_clean = df.fillna(df.mean())

# Select relevant features for clustering
features = ['X', 'Y', 'Iodine_131', 'Caesium_134', 'Caesium_137']
df_features = df_clean[features]

# Standardize the features
scaler = StandardScaler()
df_scaled = scaler.fit_transform(df_features)

Sum_of_squared_distances = []
K = range(2,8)
for num_clusters in K :
    kmeans = KMeans(n_clusters=num_clusters)
    kmeans.fit(df[["Iodine_131", "Caesium_134", "Caesium_137"]].notna())
    Sum_of_squared_distances.append(kmeans.inertia_)
plt.plot(K, Sum_of_squared_distances, "bx-")
plt.xlabel("Values of K")
plt.ylabel("Sum of squared distances/Inertia")
plt.title("Elbow Method For Optimal k")
plt.show()

# Predict on test data
y_pred = knn.predict(X_test)

# Calculate accuracy
accuracy = accuracy_score(y_test, y_pred)
print(f"Accuracy: {accuracy:.3f}")

# Predict clusters for new data
new_data = np.array([[4.00, 51.00, 0.15, 0.12, 0.25], [7.00, 47.00, 0.25, 0.18, 0.30]])
new_data_scaled = scaler.transform(new_data)
new_clusters = knn.predict(new_data_scaled)
print("Predicted clusters for new data:", new_clusters)

```

```

V.model = 'stable'
ok = OrdinaryKriging(V, min_points=2, max_points=5, mode='estimate')
xx, yy = np.mgrid[42:60:150j, 0:18:150j]
field = ok.transform(xx.flatten(), yy.flatten()).reshape(xx.shape)
s2 = ok.sigma.reshape(xx.shape)

fig, axes = plt.subplots(1, 2, figsize=(16, 8))
art = axes[0].matshow(field.T, origin='lower', cmap='plasma')
axes[0].set_title('Interpolation')
axes[0].plot(df.V, df.X, '+k')
axes[0].set_xlim((40, 54))
axes[0].set_ylim((0, 20))
plt.colorbar(art, ax=axes[0])
art = axes[1].matshow(s2.T, origin='lower', cmap='YlGn_r')
axes[1].set_title('Kriging Error')
plt.colorbar(art, ax=axes[1])
axes[1].plot(df.V, df.X, '+w')
axes[1].set_xlim((40, 54))
axes[1].set_ylim((0, 20));

def interpolate(V, ax):
    # Thanks to https://scikit-gstat.readthedocs.io/en/latest/tutorials/02_variogram_models.html

    xx, yy = np.mgrid[30:58:150j, -3:20:150j]
    ok = OrdinaryKriging(V, min_points=2, max_points=10, mode='exact')
    field = ok.transform(xx.flatten(), yy.flatten()).reshape(xx.shape)
    art = ax.matshow(field, origin='lower', cmap='plasma')
    ax.set_title('%s model' % V.model.__name__)
    plt.colorbar(art, ax=ax)
    return field

fields = []
fig, _a = plt.subplots(2, 3, figsize=(18, 12), sharex=True, sharey=True)
axes = _a.flatten()
for i, model in enumerate(('spherical', 'exponential', 'gaussian', 'matern', 'stable', 'cubic')):
    V.model = model
    fields.append(interpolate(V, axes[i]))

```

Додаток Г

ІЛЮСТРАТИВНА ЧАСТИНА

ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ГЕОСТАТИСТИЧНОГО АНАЛІЗУ
РАДІАЦІЙНОГО ЗАБРУДЖЕННЯ В ЄВРОПІ

Нормоконтроль: к.т.н., доцент


Сергій ЖУКОВ

« » 2024 р.

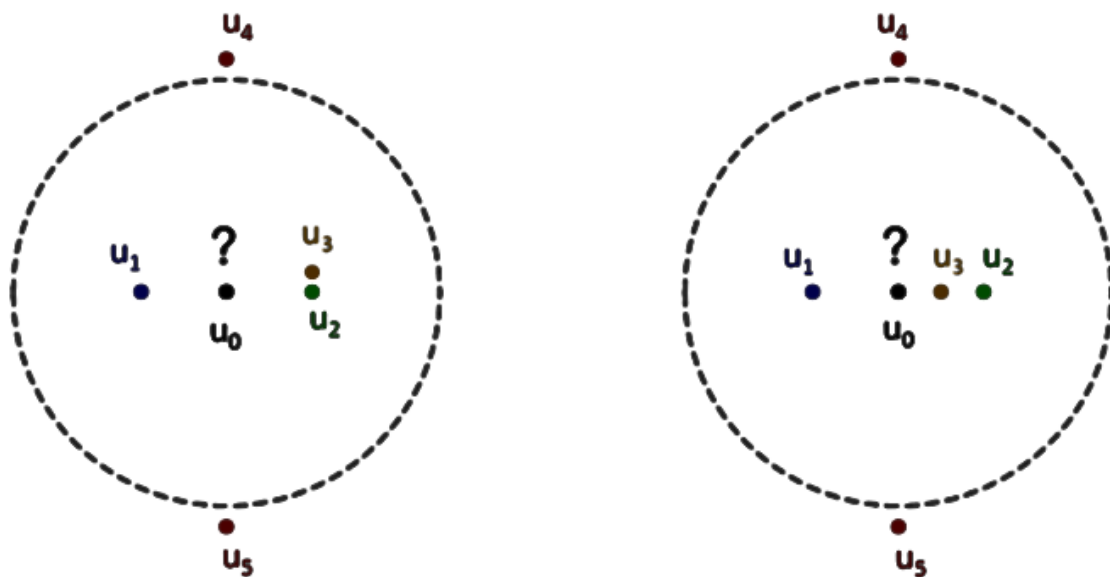


Рисунок Г.1 – Спрощений приклад кригінг-інтерполяції

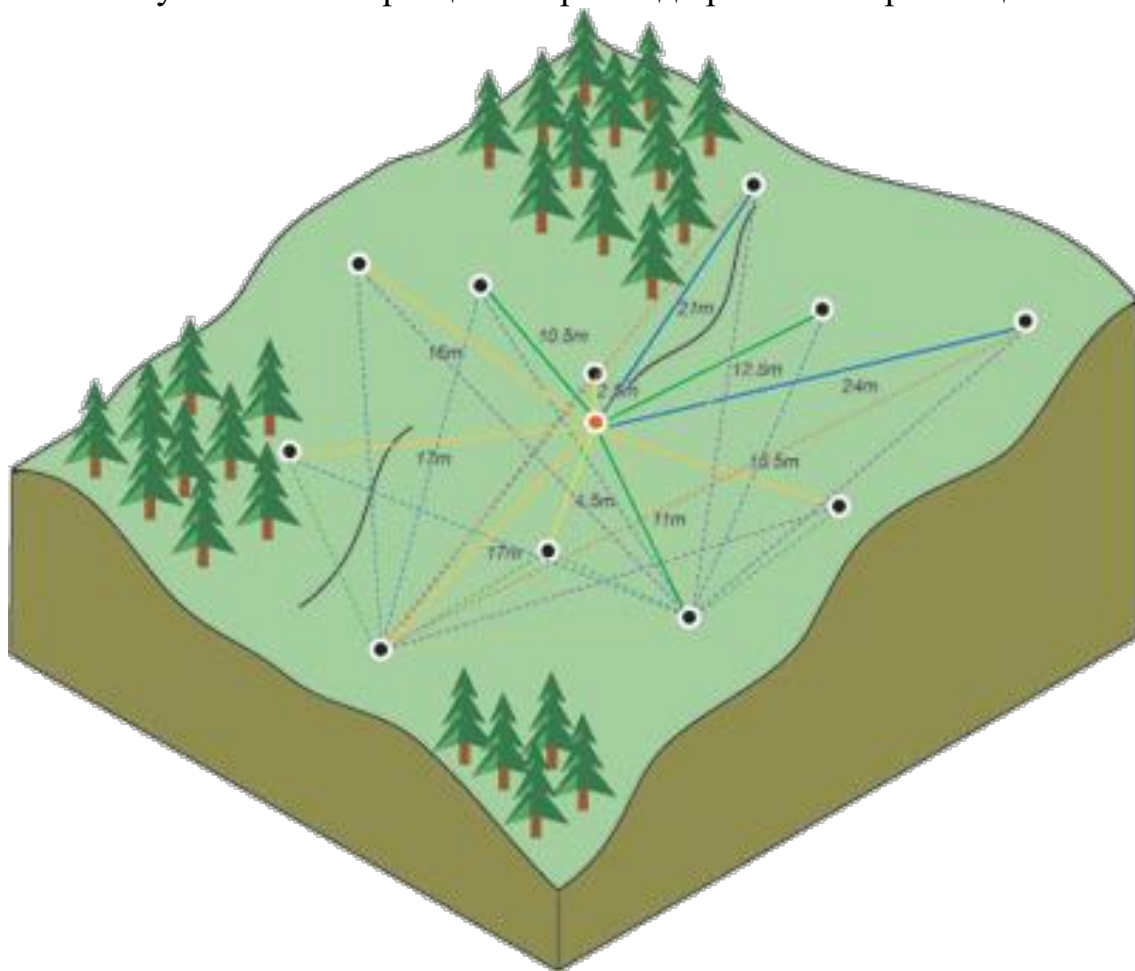


Рисунок Г.2 - Метод інтерполяції Кригінг

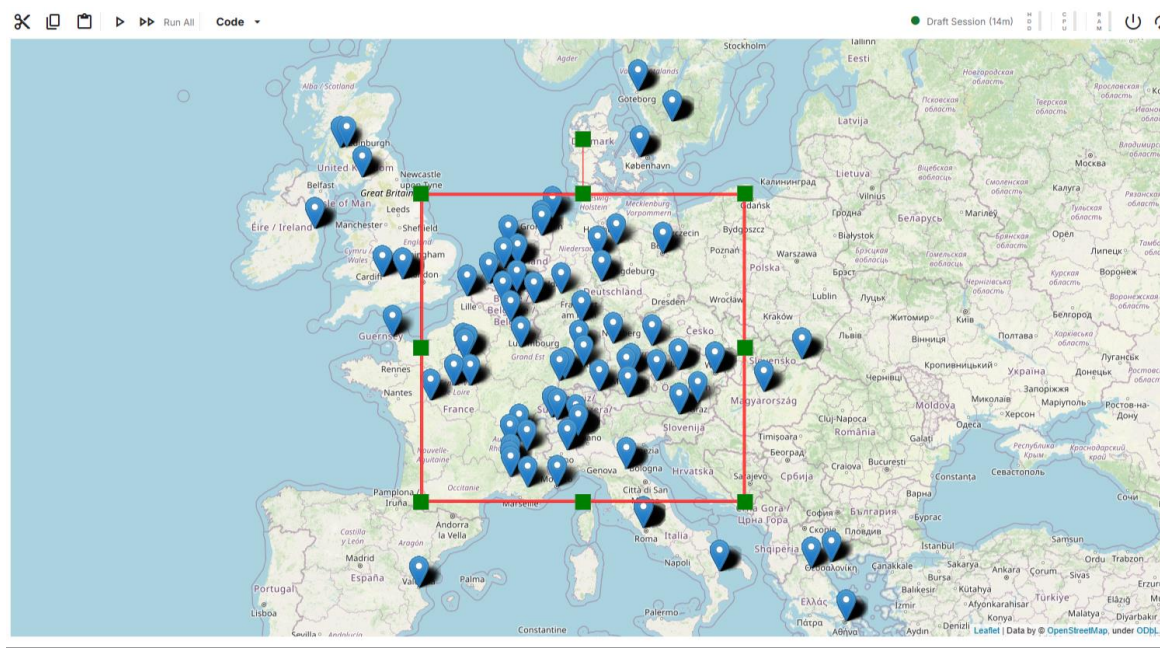


Рисунок Г.3 – Картографічне відображення вхідних даних

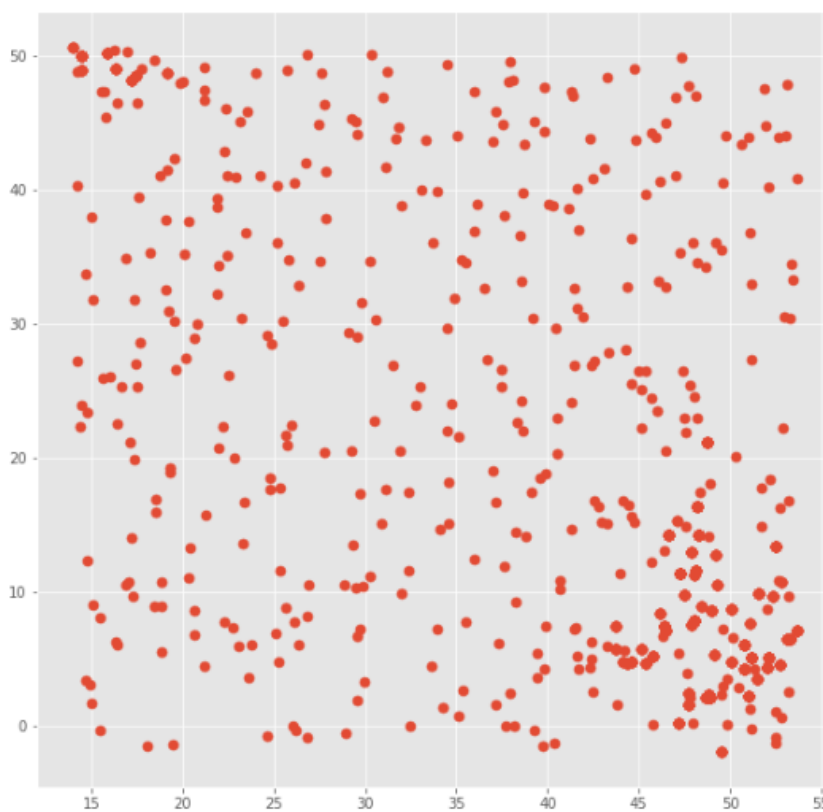


Рисунок Г.4 – Діаграма розсіювання даних після передбачення допоміжних даних

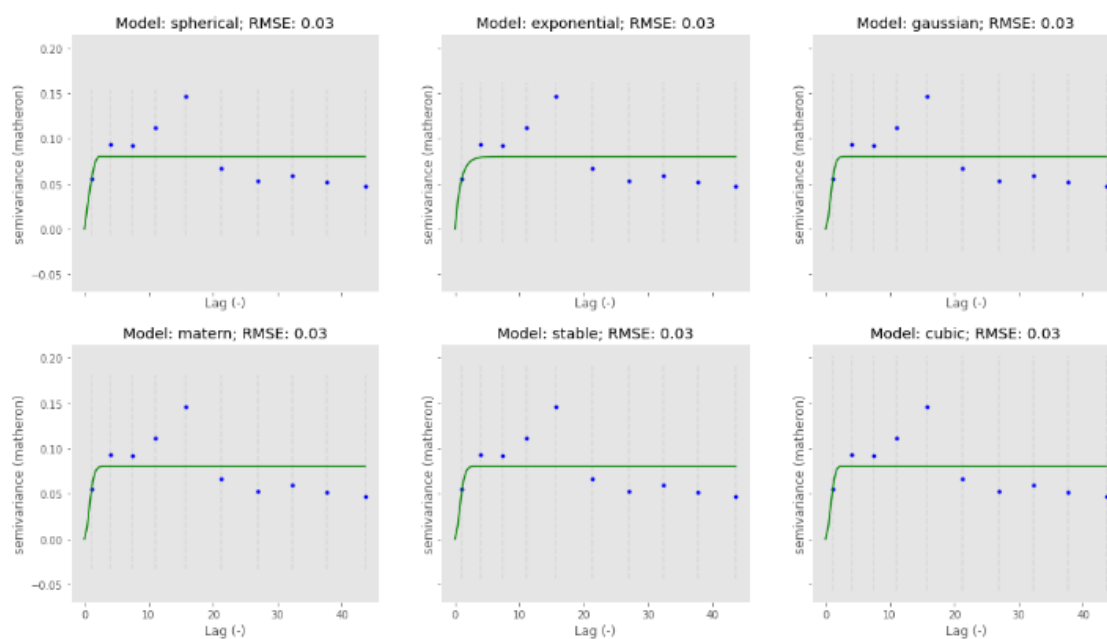


Рисунок Г.5 – Порівняння основних моделей побудови варіограм

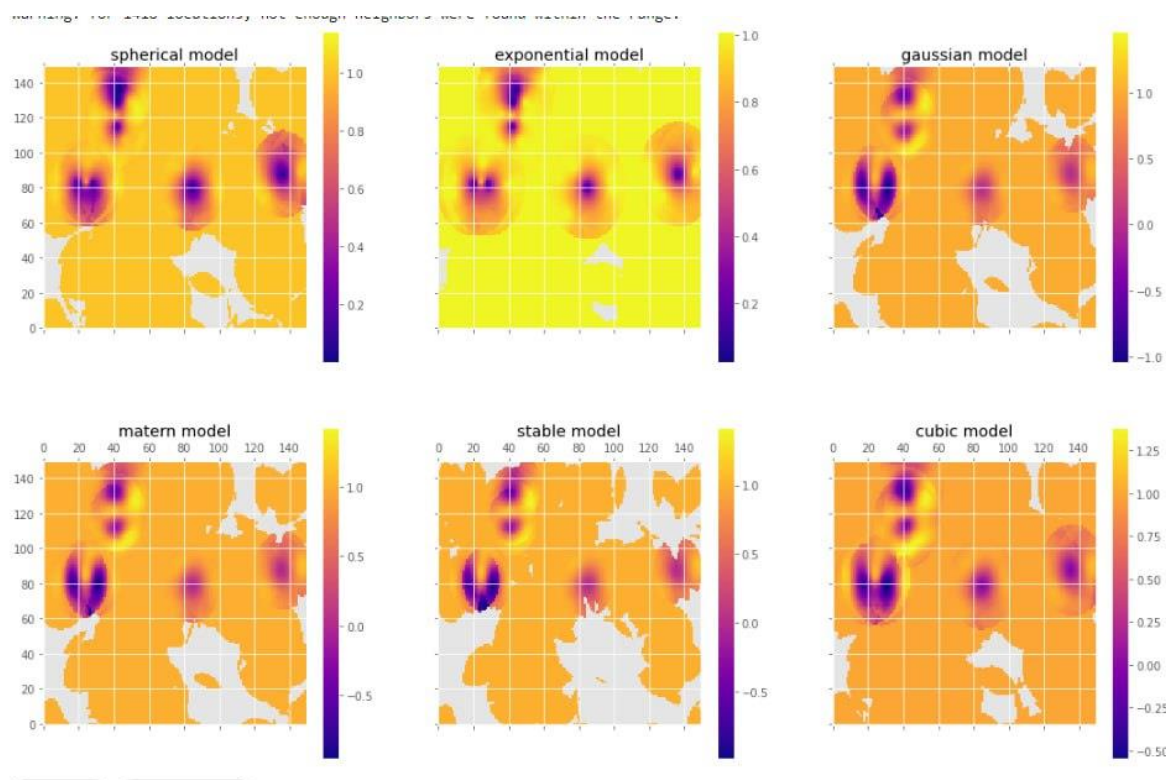


Рисунок Г.6 – Результат геостатистичного аналізу основними популярними моделями

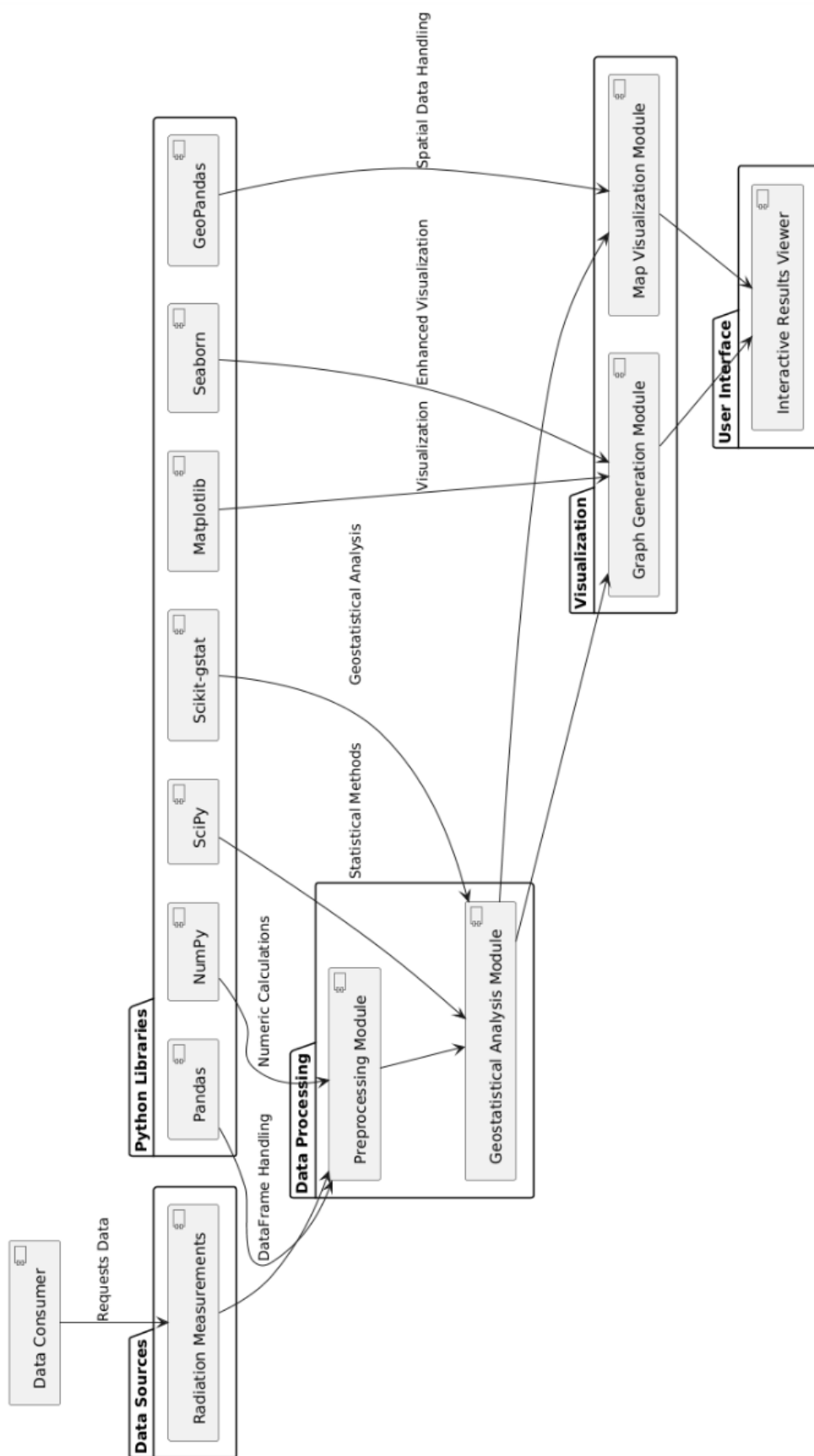


Рисунок Г.7 - UML Component Diagram