

## ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ АНАЛІЗУ СТАНУ ПОВЕРХНЕВИХ ВОД ВІННИЦЬКОЇ ОБЛАСТІ ЗА 2003–2024 РОКИ

<sup>1</sup>Вінницький національний технічний університет

### Анотація

У даній роботі розглянуто етап проектування та реалізації реляційної бази даних у середовищі MySQL для аналізу стану поверхневих вод Вінницької області. Процес розробки охоплює створення концептуальної ER-моделі та її трансформацію у фізичну структуру, що складається з пов'язаних таблиць для збереження фізико-хімічних показників та характеристик об'єктів моніторингу. Описано процес імпорту до СУБД консолідованого у середовищі MS Excel масиву даних за 2003–2024 роки. Особливу увагу приділено застосуванню SQL-інструментарію для комплексної фільтрації даних. На основі сформованих вибірок проведено розвідувальний аналіз (EDA) та побудовано прогностичні моделі концентрації амоній-іонів з використанням алгоритмів Prophet та ARIMA, що дозволило оцінити динаміку якості води та точність короткострокового прогнозування.

**Ключові слова:** MySQL, база даних, ER-модель, моніторинг вод, фільтрація даних, SQL, Вінницька область.

**Abstract** This work examines the design and implementation stage of a relational database in the MySQL environment for analyzing the state of surface waters in the Vinnytsia region. The development process covers the creation of a conceptual ER model and its transformation into a physical structure consisting of linked tables for storing physicochemical parameters and monitoring object characteristics. The process of importing a consolidated dataset for the years 2003–2024 from MS Excel into the DBMS is described. Special attention was paid to the use of SQL tools for complex data filtering. Based on the generated samples, exploratory analysis (EDA) was conducted and predictive models of ammonium ion concentration were built using Prophet and ARIMA algorithms, which allowed assessing the dynamics of water quality and the accuracy of short-term forecasting.

**Keywords:** MySQL, database, ER model, water monitoring, data filtering, SQL, Vinnytsia region.

### Вступ

Сучасний стан екологічної безпеки вимагає впровадження ефективних інструментів для оперативного аналізу та контролю якості водних ресурсів [1]. Однією з актуальних проблем у цій сфері є розпорошеність та великий обсяг накопичених даних моніторингу, що ускладнює процес виявлення довгострокових трендів забруднення та ідентифікацію аномальних викидів [2-4]. Традиційні методи збереження інформації у вигляді таблиць не завжди забезпечують необхідну швидкість обробки, цілісність та можливість зручної фільтрації даних за територіальною або часовою ознакою.

Для вирішення цієї проблеми пропонується розробка інформаційної технології, що базується на використанні реляційних баз даних. Це дозволить централізувати багаторічні спостереження, автоматизувати процеси вибірки даних та забезпечити надійний фундамент для подальшого аналізу екологічного стану регіону.

Метою роботи є розробка та реалізація бази даних у середовищі MySQL [5] для аналізу стану поверхневих вод Вінницької області. Система дозволить консолідувати дані моніторингу за 2003–2024 роки, проводити їх швидко фільтрацію та підготовку часових рядів для подальшого моделювання. Основними завданнями дослідження є проектування концептуальної ER-моделі бази даних, розробка її фізичної структури, організація імпорту консолідованих даних та реалізація програмних запитів для формування цільової вибірки, проведення розвідувального аналізу та побудова прогностичних моделей для оцінки екологічного стану водних об'єктів.

### Результати дослідження

У ході дослідження було розроблено архітектуру та технічне наповнення реляційної бази даних, що є ядром інформаційної технології аналізу стану водних ресурсів. Структура бази даних у

MySQL була спроектована таким чином, щоб забезпечити швидку обробку великих масивів інформації та цілісність зв'язків між об'єктами моніторингу. Реалізована система забезпечує виконання таких ключових функцій:

- **Централізоване зберігання даних** — консолідація багаторічних спостережень за хімічними показниками (амоній, фосфати тощо) в єдиному структурованому середовищі.
- **Територіальна прив'язка** — завдяки використанню таблиць monitoringposts та localities, кожен запис про стан води автоматично корелюється з конкретною локацією у Вінницькій області.
- **Гнучка фільтрація** — програмна реалізація SQL-запитів дозволяє миттєво виокремлювати дані за певними часовими інтервалами (2003–2024 рр.) або за конкретними водними об'єктами.
- **Підготовка даних для аналізу** — автоматизоване об'єднання таблиць через оператори JOIN створює готові набори даних (datasets) для подальшої візуалізації та статистичної обробки на Python.

Технічна реалізація бази даних передбачає використання середовища MySQL Workbench для адміністрування та написання SQL-скриптів. Наповнення бази здійснювалося шляхом імпорту попередньо консолідованого в MS Excel масиву даних, що дозволило уникнути втрати інформації при переході від табличних форм до реляційної моделі. Розроблена фізична модель БД, а також створені SQL-скрипти для генерації таблиць (CREATE TABLE) та маніпулювання даними (INSERT, SELECT), стали надійним фундаментом для повноцінного функціонування інформаційної технології аналізу водних об'єктів.

На рисунках 1–2 представлено візуалізацію архітектурних рішень, а саме концептуальну ER-модель та фізичну структуру реляційної бази даних, які стали основою для реалізації інформаційної технології аналізу стану поверхневих вод.

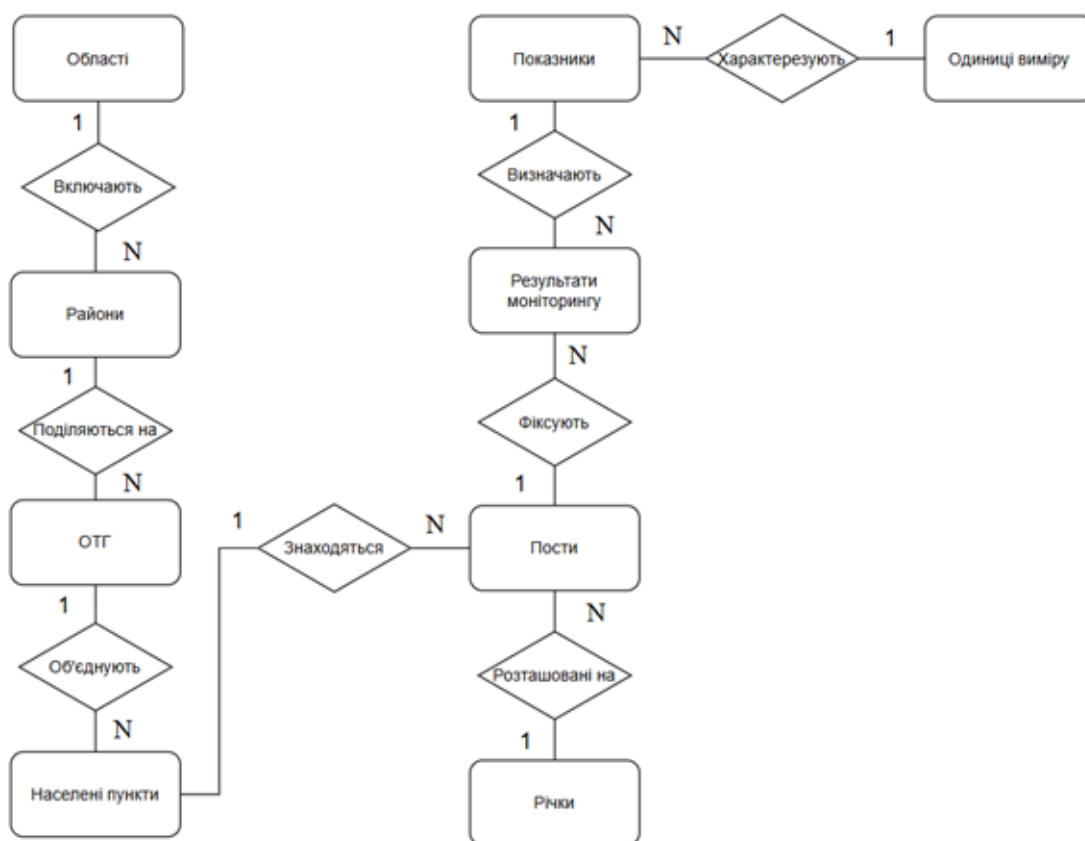


Рисунок 1 – ER-модель бази даних моніторингу поверхневих вод

На даному рисунку зображено логічну структуру даних та зв'язки типу «один-до-багатьох» (1:N), що забезпечують цілісність інформаційної системи:

1. **Адміністративна ієрархія:** логічний ланцюжок від «Області» через «Райони» та «ОТГ» до конкретних «Населених пунктів», що дозволяє проводити точну просторову фільтрацію.

2. **Блок моніторингу:** сутність «Пости» є центральною ланкою, яка пов'язана з «Річками» (місце розташування) та «Результатами моніторингу».
3. **Блок показників:** сутність «Показники» пов'язана з «Одиницями виміру» та визначає параметри кожного конкретного результату заміру.

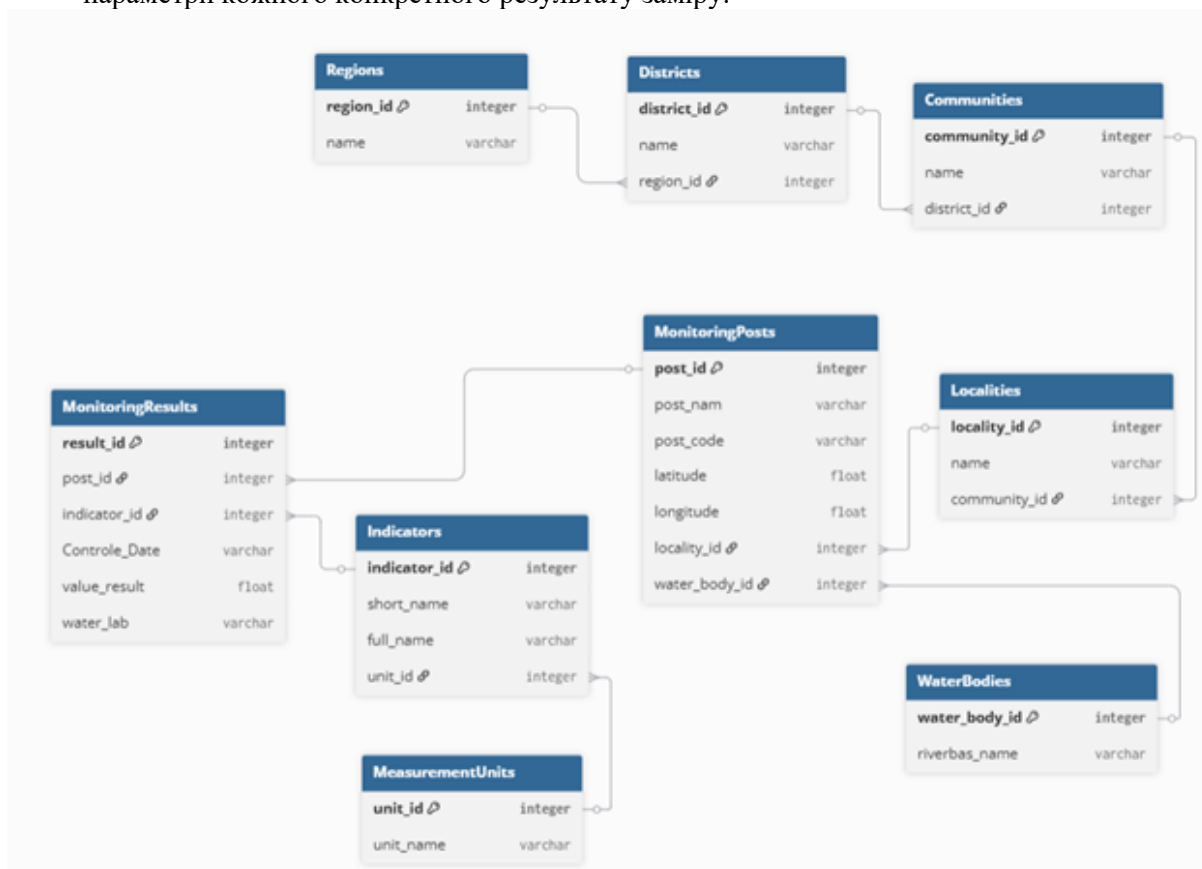


Рисунок 2 – Фізична модель бази даних у середовищі MySQL

На рисунку представлена детальна структура таблиць із зазначенням типів даних (integer, varchar, float) та встановлених зв'язків через первинні та зовнішні ключі:

1. **Таблиця MonitoringResults:** збирає дані замірів, включаючи дату контролю (Controle\_Date), значення концентрації (value\_result) та посилання на лабораторію.
2. **Таблиця MonitoringPosts:** містить технічні дані постів, зокрема точні географічні координати (latitude, longitude) для картографічної візуалізації.
3. **Довідникові таблиці:** Regions, Districts, Communities та Localities утворюють нормалізовану структуру для зберігання адміністративних даних Вінницької області.
4. **Таблиці Indicators та MeasurementUnits:** забезпечують стандартизацію найменувань хімічних показників та їх одиниць вимірювання (наприклад, мг/дм<sup>3</sup>).

Розроблена структура дозволяє ефективно виконувати складні SQL-запити з використанням операторів для об'єднання екологічної та географічної інформації в межах єдиної інформаційної технології.

Наступним етапом дослідження став розвідувальний аналіз (EDA) сформованого масиву даних. Для виявлення внутрішніх взаємозв'язків побудовано кореляційну матрицю (рис. 3) за даними створу Сабарівського вдсх (м. Вінниця). Виявлено сильні прямі зв'язки між показниками органічного забруднення: HSK та Permanganat (0.89), а також специфічну кореляцію між Nitrat та Atrazin (0.78), що свідчить про вплив аграрного сектору на якість води.

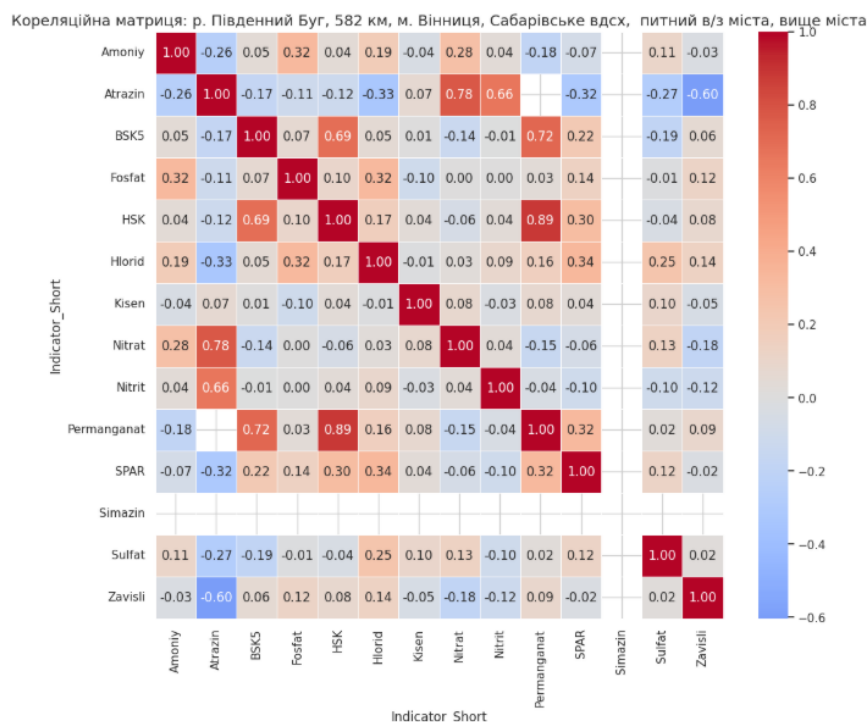


Рисунок 3 – Кореляційна матриця

Для прогнозування екологічного стану було обрано показник Амоній-іони. Порівняльний аналіз моделей показав перевагу алгоритму Prophet [6] (MAE = 0.2370) над класичною моделлю ARIMA [7] (MAE = 0.2464). Проведено тюнінг гіперпараметрів моделі Prophet (оптимізація сезонності та точок зміни тренду), що дозволило знизити помилку до 0.2263 (ефективність покращення складала 4.49%). Отримані результати підтверджують придатність розробленої технології для завдань моніторингу та оперативного прогнозування якості водних ресурсів.

### Висновки

Розроблена реляційна база даних у MySQL довела свою ефективність як інструмент консолідації та підготовки даних для інтелектуального аналізу. У ході роботи не лише реалізовано структуру зберігання, а й практично застосовано методи машинного навчання для прогнозування вмісту амоній-іонів. Встановлено, що модель Prophet після процедури тюнінгу забезпечує вищу точність прогнозу порівняно з ARIMA. Це дозволяє рекомендувати створену інформаційну технологію для використання в системах екологічного моніторингу Вінницької області з метою раннього попередження про зміну якості поверхневих вод.

Подальший розвиток системи може включати розширення функціоналу, а саме: автоматизацію процесу завантаження нових даних моніторингу через веб-інтерфейс адміністратора, інтеграцію бази даних із ГІС-системами для автоматичної візуалізації забруднень на інтерактивних картах, а також розробку процедур збереження результатів прогнозних моделей, отриманих за допомогою алгоритмів машинного навчання. Це дозволить перетворити розроблену структуру на повноцінну систему підтримки прийняття рішень у сфері охорони водних ресурсів регіону.

### СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Мокін В. Б., Крижановський Є. М. Математичне моделювання динаміки якості річкових вод. – Вінниця: ВНТУ, 2023. – 160 с.
2. Державне агентство водних ресурсів України. Моніторинг поверхневих вод. – 2025. [Електронний ресурс]. – Режим доступу: <https://www.davr.gov.ua/monitoring-poverhnevih-vod1>
3. Басейнове управління водних ресурсів р. Південний Буг. Моніторинг стану якості поверхневих вод. – 2025. [Електронний ресурс]. – Режим доступу: <https://buvrpb.davr.gov.ua/novyny/monitorynh-stanu-iakosti-poverkhnevykh>
4. Моніторинг та екологічна оцінка водних ресурсів України: інтерактивна система Держводагентства. – 2024. [Електронний ресурс]. – Режим доступу: <http://monitoring.davr.gov.ua>

5. MySQL — Офіційна документація. – 2024. [Електронний ресурс]. – Режим доступу: <https://dev.mysql.com/doc/>
6. Prophet: Forecasting at scale. Facebook Open Source. – 2024. [Електронний ресурс]. – Режим доступу: <https://facebook.github.io/prophet/>
7. Box G. E. P., Jenkins G. M. Time Series Analysis: Forecasting and Control. – 2015. – 568 p.

**Анастасія Максимівна Герасимович** – студентка групи 2ICT-226, факультет інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, Вінниця, e-mail: [herasymovcyh.a.m@gmail.com](mailto:herasymovcyh.a.m@gmail.com)

**Євгеній Миколайович Крижановський** – канд. техн. наук, доцент кафедри системного аналізу та інформаційних технологій, Вінницький національний технічний університет, Вінниця, e-mail: [kruzhan@gmail.com](mailto:kruzhan@gmail.com)

**Ігор Миколайович Штельмах** – канд. техн. наук, асистент кафедри системного аналізу та інформаційних технологій, Вінницький національний технічний університет, Вінниця, e-mail: [igor.shtelmakh@vntu.edu.ua](mailto:igor.shtelmakh@vntu.edu.ua).

**Herasyimovych Anastasiia M.** – student of Faculty of Intelligent Information Technology and Automation, 2IST-21b, Vinnytsia National Technical University, Vinnytsia, e-mail [annashport11@gmail.com](mailto:annashport11@gmail.com)

**Kryzhanovsky, Evgeniy M.** - candidate of technical sciences, associate professor of the Department of System Analysis and Information Technologies, Vinnytsia National Technical University, Vinnytsia, e-mail: [kruzhan@gmail.com](mailto:kruzhan@gmail.com)

**Shtelmah, Igor M.** – candidate of technical sciences, assistant professor of the Department of System Analysis and Information Technologies, Vinnytsia National Technical University, Vinnytsia, e-mail: [igor.shtelmakh@vntu.edu.ua](mailto:igor.shtelmakh@vntu.edu.ua).