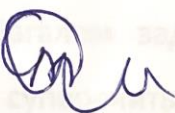

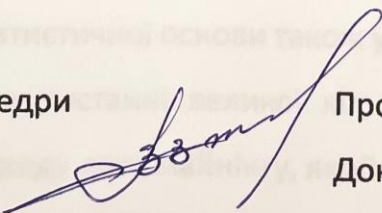


Міністерство освіти і науки України
Київський національний університет імені Тараса Шевченка
ННІ «Інститут геології»
Кафедра Геоінформатики

КВАЛІФІКАЦІЙНА РОБОТА МАГІСТРА
спеціальність 103 – Науки про Землю
освітня програма «Геоінформатика»
спеціалізований блок «Дистанційне зондування»

ТЕМА: «Проектування системи прийняття рішень для задач електророзвідки»

Виконав		студент 2-го курсу ОС «Магістр» кафедри Геоінформатика
Науковий керівник		Доцент Кандидат геологічних наук Віршило Іван Вікторович
Робота рекомендується до захисту (протокол № 8 засідання кафедри геоінформатика від 15 травня 2023р.)		
Завідувач кафедри		Професор Доктор технічних наук Зацерковний Віталій Іванович

Київ – 2023

РЕФЕРАТ

Магістерської кваліфікаційної роботи

Столпакова Андрія Юрійовича

На тему: «Проектування системи прийняття рішень для задач електророзвідки»

Спеціальність **103 Науки про Землю**

Робота на 78 аркушів складається з 3 розділів та 8 підрозділів й містить 5 таблиць, 24 рисунків має 2 додатки та 11 підрозділів в додатку А. При підготовці роботи використовувалися матеріали 49 джерел.

Актуальність: Задачі геологічної інтерпретації геофізичних даних містять високу ступінь невизначеності і виконуються на основі знань та досвіду інтерпретаторів. Загалом задача зводиться до побудови геоелектричного розрізу, що не суперечить спостереженим даним електророзвідки та петрофізичним характеристикам геологічних товщ. Такі задачі відносяться до класу некоректно поставлених, а нескінчена множина їх аналітичних розв'язків містить багато наборів, що не мають геологічного сенсу. Тому подібні задачі на поточному рівні розвитку галузі прийнято вирішувати через геолого-геофізичне моделювання. Нові підходи можуть використовувати стохастичні методи. Але оскільки немає уніфікованих наборів даних для різних геологічних ситуацій та методик проведення дослідження, отримання достовірно статистичної основи також утруднене. Тому сьогодні постають нові виклики у використанні великої кількості наявних різномірних даних. Це відповідає підходу даних майнінгу, який активно розвивається в різних галузях, в тому числі і в геології. Через це застосування підходу великих даних Big data, в тому числі його складових частин: підготовки та зберігання даних, data mining та візуалізація є дуже актуальним на сьогоднішній день.

Мета роботи: Спроекувати систему підтримки прийняття рішень, на основі підходу великого масиву даних, для вирішення задачі прогнозування електричного опору у верхній частині геологічного розрізу.

Розв'язувані в роботі задачі:

- 1) Проаналізувати, які фактори впливають на електропровідність гірських порід.
- 2) Проаналізувати пряму та обернену задачу електророзвідки.
- 3) Проаналізувати типи електричних установок.
- 4) Спроекувати сховище даних для датасетів із характеристик, пов'язаних з задачею вимірювання електричного опору геологічного середовища.
- 5) Продемонструвати можливості використання для комплексного аналізу впливу різних факторів на виміряний опір.

Основний зміст: Застосування підходу розробки системи підтримки прийняття рішень для даних електротомографії. Розроблення схеми для СППР на основі даних, які можуть впливати на виміряні значення опору.

Наукові результати: Застосовано підхід Big data, а саме його складової зберігання багатовимірних даних у data warehouse для геоелектричних даних. В рамках даного підходу було застосовано ROLAP для вимірювань електротомографії. В межах даного підходу було розроблено багатовимірну базу даних типу «сніжинка» в ній було виділено 10 розмірностей.

Наукова новизна: Створено багатовимірну модель та спроектовано та реалізовано схему сховища для геоелектричних даних для системи підтримки прийняття рішень в задачах прогнозування геоелектричного опору шаруватих середовищ.

Практичне значення: Створення інструменту для можливої комплексної інтерпретації даних геоелектрики з іншими факторами, що впливають на електропровідність геологічного середовища. Даних підхід можна застосувати для подальшого аналізу методами машинного навчання.

Анотація

В даній роботі була спроектована система підтримки прийняття рішень по типу ROLAP. В рамках, якої було реалізовано багатовимірну базу даних. Також було продемонстровано виділення розмірностей в такій системи для задач електророзвідки.

Вхідними матеріалами для даної роботи слугували дані польових вимірювань методом ERT, сіткою диполь-диполь. Цих даних було виявилось замало для реалізації такої задачі. В рамках даної роботи було створено базу даних за допомогою системи керування базами даних PostgreSQL. В ній було виділено 10 розмірностей таких, як: площа, тип вимірювання (array number), хронологічна шкала, підземні води, вимірювальний пристрій, оператор, погода, а також літологія та її ієрархічну розмірність літологічний підтип.

Було зроблено висновки про необхідність доповнення даної бази даних іншими системами зокрема на основі штучних мереж, що в свою чергу дасть суттєве збільшення точності прогнозування на основі даних електророзвідки.

Ключові слова: Багатовимірна база даних, система підтримки прийняття рішень, електрозеристивна томографія

Stolpakov Andrii

Designing a decision-making system for electrical exploration tasks

Abstract

In this paper, a ROLAP-type decision support system was designed. Within the framework of which a multidimensional database was implemented. We have also demonstrated the selection of dimensions in such a system for electrical exploration tasks.

The input materials for this work were field measurements using the ERT method and a dipole-dipole grid. This data was not enough to solve this problem. As part of this work, a database was created using the PostgreSQL database management

system. It was divided into 9 dimensions, such as: area, type of measurement (array number), chronological scale, groundwater, measuring device, operator, weather, as well as lithology and its hierarchical dimension lithological subtype.

It was concluded that it is necessary to supplement this database with other systems, in particular, those based on artificial networks, which in turn will significantly increase the accuracy of forecasting based on electrical survey data.

Key words multidimensional database, decision support system, electrical resistivity tomography

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ.....	7
ВСТУП	8
1. Огляд літератури	12
1.1 Опис систем підтримки прийняття рішень	12
1.2 Огляд стану застосування СППР та баз даних в геології	22
2. Багатовимірні бази даних	38
2.1 Багатовимірна аналітична обробка, як вид багатовимірної бази даних	47
2.2 Реляційна аналітична обробка, як вид багатовимірної бази даних.....	50
2.2.1 Опис схеми ROLAP типу «Зірка»	52
2.2.2. Опис схеми ROLAP типу «Сніжинка».....	54
2.3 Гібридна аналітична обробка, як вид багатовимірної бази даних.....	57
3. ПРАКТИЧНА ЧАСТИНА.....	61
3.1 Огляд методу електрорезистивна томографії.....	63
3.2 Розробка багатовимірної бази даних.....	68
3.3 Заповнення таблиць	78
ВИСНОВКИ	86
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ:	88
Додаток А.....	95
Додаток А.1	95
Додаток А.2	95
Додаток А.3	95
Додаток А.4	96
Додаток А.5	96
Додаток А.6	97
Додаток А.7	97
Додаток А.8	98
Додаток А.9	98
Додаток А.10	98
Додаток А.11	99
Додаток Б.....	102

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

ERT – з англійської мови Electrical resistivity tomography - Електрорезистивна томографія

HOLAP - з англійської мови Hybrid Online Analytical Processing – гібридна онлайн-аналітична обробка.

IP – з англійської мови Inducted polarization - Індукована поляризація

MOLAP - з англійської мови Multidimensional Online Analytical Processing - Багатовимірна онлайн-аналітична обробка.

OLAP – з англійської мови Online Analytical Processing - онлайн-аналітична обробка.

RDBMS – з англійської мови Relation database management system - реляційна система керування базою даних.

ROLAP - з англійської мови Relational Online Analytical Processing - реляційна онлайн-аналітична обробка.

SQL – з англійської мови Structured Query Language - Мова структурованих запитів.

БД – база даних.

СКБД – Системи керування базами даних.

ВСТУП

Дані електророзвідки як явище є дуже не самостійним їх важко розглядати, окремо від інших геофізичних досліджень. Взагалі всі геофізичні методи по своїй природі спрямовані на розв'язання прямої і зворотної задачі це в свою чергу термін математичної фізики, що відіграє службову роль у розвідувальній геофізиці. Через це виділяють пряму та не пряму задачі геофізики.

Пряма задача полягає визначенні аномалій (їхньої величини, форми і положення) різними методами. Наприклад сейсмікою, акустичним зондування чи як в даній роботі електрорезистивною томографією. Все залежить від задачі, та можливостей. Це все робить за відомим розташуванням об'єктів, що мають задані фізичні характеристики.

Обернена задача це - визначення за спостережуваними аномаліями розташування об'єктів, що викликають аномалії, їхніх геометричних і фізичних параметрів. У складних випадках отримують наближене розв'язання зворотної задачі або кілька розв'язань, частину яких відкидають у разі введення додаткових обмежень. У зв'язку з тим, що дана задача, для електророзвідки немає аналітичного рішення через її неправильне формулювання. Через це, існують різного роду наближені методи, які в тій чи іншій мірі дозволяють отримати знання про геології досліджуваного регіону. У зв'язку з цим актуальним є підвищення ефективності розв'язання оберненої задачі електророзвідки для встановлення геологічного розрізу.

Застосування інших підходів, як аналізу так і зберігання даних електророзвідки є дуже мало дослідженим питанням. Для великих масивів даних існують свої методи, наприклад Big data, вони для геофізики в цілому рідко застосовуються, а для електрики взагалі можна сказати, що такий підхід не застосовується.

Системи прийняття рішень, як інтерпретації Big data відіграють важливу роль в організаціях із різних галузей, та дають можливість прийняти обґрунтовані рішення. Ці системи значною мірою покладаються на бази даних, які слугують центральним сховищем інформації, що використовується для аналізу.

Використовуючи передові технології, такі, як машинне навчання і штучний інтелект, системи прийняття рішень можуть швидко обробляти величезні обсяги даних і надавати інформацію, яку неможливо отримати вручну, або для її отримання знадобилося дуже багато часу. Це дає змогу людям, які приймають рішення, виявляти тенденції, закономірності та взаємозв'язки, які можуть бути не очевидними, що дає їм змогу приймати більш обґрунтовані рішення та швидко реагувати на мінливі ринкові умови. Також машинне навчання може допомогти знайти не явні закономірності, між даними, які можуть впливати на вимірний опір.

Важливість систем прийняття рішень та їх головної складової баз даних, неможливо переоцінити. Вони мають вирішальне значення для того, щоб бізнес залишався конкурентоспроможним і приймав обґрунтовані рішення, які можуть сприяти зростанню та успіху. Тому організації повинні інвестувати в розвиток і підтримку цих систем і баз даних, щоб вони залишалися надійними, безпечними та ефективними упродовж тривалого часу.

Актуальність теми: Задачі геологічної інтерпретації геофізичних даних містять високу ступінь невизначеності і виконуються на основі знань та досвіду інтерпретаторів. Загалом задача зводиться до побудови геоелектричного розрізу, що не суперечить спостереженням даних електророзвідки та петрофізичним характеристикам геологічних товщ. Такі задачі відносяться до класу некоректно поставлених, а нескінченна множина їх аналітичних розв'язків містить багато наборів, що не мають геологічного сенсу. Тому подібні задачі на поточному рівні розвитку галузі прийнято вирішувати через

геолого-геофізичне моделювання. Нові підходи можуть використовувати стохастичні методи. Але оскільки немає уніфікованих наборів даних для різних геологічних ситуацій та методик проведення дослідження, отримання достовірно статистичної основи також утруднене. Тому сьогодні постають нові виклики у використанні великої кількості наявних різномірних даних. Це відповідає підходу даних майнінгу, який активно розвивається в різних галузях, в тому числі і в геології. Через це застосування підходу великих даних Big data, в тому числі його складових частин: підготовки та зберігання даних, data mining та візуалізація є дуже актуальним на сьогоднішній день.

Мета роботи: Спроекувати систему підтримки прийняття рішень, на основі підходу великого масиву даних, для вирішення задачі прогнозування електричного опору у верхній частині геологічного розрізу.

Об'єкт: Електричний опір Землі виміряний методом електротомографії.

Предмет досліджень: Виділення всіх можливих чинників, які впливають на електричний опір, який було виміряно методом електротомографії.

Інформація про методологічні засади та методи дослідження: Під час написання даної роботи були використані наступні методи досліджень: опис, аналіз та синтез інформації.

Завдання: 1) Проаналізувати, які фактори впливають на електропровідність гірських порід.

2) Проаналізувати пряму та обернену задачу електророзвідки.

3) Проаналізувати типи електричних установок.

4) Спроекувати сховище даних для датасетів із характеристик, пов'язаних з задачею вимірювання електричного опору геологічного середовища.

5) Продемонструвати можливості використання для комплексного аналізу впливу різних факторів на виміряний опір.

Наукова новизна: використання підходу Big data, а саме його складової data warehouse для геоелектричних даних.

Практичне значення: Створення інструменту для можливої комплексної інтерпретації даних геоелектрики з іншими факторами, що впливають на електропровідність геологічного середовища. Даних підхід можна застосувати для подальшого аналізу методами машинного навчання

Особистий внесок здобувача: Виділення розмірностей для багатовимірної бази даних та її розроблення.

Апробації: З тезами на дану тематику виступав на XIII Всеукраїнській молодіжній науковій конференції-школі "Сучасні проблеми наук про Землю", в секції: «Геоінформатика та сучасні інформаційні технології в науках про Землю;»

Публікації:

Джерела походження вхідних даних: *неопубліковані лабораторні дані.*

Загальні відомості про структуру і обсяг роботи: робота розміщена на 75 аркушах та містить 3 розділи.

1. Огляд літератури

У сучасному бізнес-середовищі організації стикаються з постійно зростаючим обсягом даних, яким необхідно управляти та аналізувати для прийняття обґрунтованих рішень. Саме тут у гру вступають системи прийняття рішень та їх ядро бази даних. Система підтримки прийняття рішень (СППР) — це комп'ютерна інформаційна система, яка підтримує діяльність із прийняття ділових або організаційних рішень, що зазвичай полягає в ранжуванні, сортуванні або виборі з-поміж альтернатив. Такі системи обслуговують управлінський, операційний та плановий рівні організації (зазвичай це ланки середнього та вищого менеджменту) і допомагають людям приймати рішення щодо проблем, які можуть швидко змінюватися і які важко визначити заздалегідь (тобто неструктуровані та напівструктуровані проблеми прийняття рішень) (Pal, 2019).

Ці системи використовують різні методи, такі, як інтелектуальний аналіз даних, прогнозне моделювання та алгоритми машинного навчання, для виявлення закономірностей і висновків із великих наборів даних. Бази даних у таких системах використовуються для зберігання та організації цих даних, що полегшує доступ до них та їх аналіз.

Розуміння структури, типів та прикладного застосування кожної системи є дуже важливим, оскільки потрібно чітко розуміти переваги та недоліки таких систем.

1.1 Опис систем підтримки прийняття рішень

Знання структури й типу системи підтримки прийняття рішень є важливим, оскільки воно дає змогу організаціям вибрати правильну систему для своїх конкретних потреб і зрозуміти, як її ефективно використовувати.

Завдяки цьому, організації можуть підвищити якість та узгодженість своїх рішень, покращити комунікацію та співпрацю, а також впоратися з інформаційним перевантаженням. Тепер детально розглянемо типи та структури таких систем.

Загалом, вибір типу системи підтримки прийняття рішень залежить від характеру проблеми прийняття рішень і ресурсів, якими володіє організація. Вибравши відповідну систему підтримки прийняття рішень або комбінацію систем, організації можуть покращити процес прийняття рішень і, зрештою, досягти кращих результатів. Загалом виділяють чотири типи систем підтримки прийняття рішень:

1)СППР, орієнтовані на комунікацію: цей тип сприяє комунікації та співпраці між особами, які приймають рішення. Вона надає інструменти й технології, які даю змогу особам, що приймають рішення, спілкуватися один з одним, обмінюватися інформацією і знаннями та спільно приймати рішення. Система на основі комунікацій, особливо, корисна для прийняття рішень у географічно розподілених організаціях, де особи, які приймають рішення, знаходяться в різних місцях (*Sharda, Delen, & Turban, 2014*).

Прикладом СППР, орієнтованої на комунікацію, є програма Cisco Webex — система веб-конференцій, яка дозволяє користувачам організувати та відвідувати віртуальні зустрічі, вебінари та тренінги. Система надає такі функції, як спільний доступ до екрану, дошка, чат і запис, які можуть допомогти учасникам ефективно співпрацювати та спілкуватися (*офіційний сайт Webex*). Дана програма, може, використовуватися організаціями різних розмірів і галузей, включаючи охорону здоров'я, освіту і фінанси, для полегшення проведення віртуальних зустрічей і поліпшення процесу прийняття рішень.

2)СППР на основі даних: такий тип системи зосереджується на аналізі великих обсягів даних. Він допомагає особам, які приймають рішення, аналізувати дані з різних джерел, виявляти закономірності й тенденції та

приймати обґрунтовані рішення на основі аналізу. Система на основі даних використовує різні методи, такі, як інтелектуальний аналіз даних, предиктивна аналітика та машинне навчання для аналізу даних (*Sharda, Delen, & Turban, 2014*).

Одним із прикладів СППР на основі даних є система управління взаємовідносинами з клієнтами (customer relationship management- CRM), яка використовує дані про клієнтів для покращення маркетингових та продажних рішень. Наприклад, менеджер із маркетингу, може, використовувати таку систему, яка була створена на основі даних для аналізу інформації про клієнтів і містить таку інформацію, як історія покупок і поведінка в Інтернеті, щоб виявити тенденції та закономірності. Система, може, надавати інформацію та рекомендації, які допоможуть особам, що приймають рішення, оптимізувати свої маркетингові кампанії та покращити залучення і утримання клієнтів.

Прикладом технічної реалізації цієї системи є програмне забезпечення Salesforce CRM — хмарна CRM-система, яка допомагає організаціям керувати продажами, маркетингом та обслуговуванням клієнтів. Програмне забезпечення використовує інформацію про клієнтів, такі, як інформація про потенційних клієнтів, історія облікових записів і взаємодія в соціальних мережах, для надання аналітики та рекомендацій, які можуть допомогти організаціям покращити взаємодію з клієнтами та ефективність продажів (*офіційний сайт CRM Salesforce*). Воно надає такі функції, як візуалізація даних, предиктивна аналітика та штучний інтелект, які можуть допомогти особам, що приймають рішення, приймати рішення на основі інформації і покращувати результати бізнесу.

3) СППР на основі моделей: цей тип системи зосереджується на використанні моделей та симуляцій для підтримки прийняття рішень. Він допомагає особам, які приймають рішення, створювати моделі, що представляють собою сценарії реального світу, та імітувати вплив різних

рішень на ці сценарії. СППР на основі моделей, особливо, корисна для прийняття складних рішень у таких сферах, як фінанси, операції та логістика (*Sharda, Delen, & Turban, 2014*).

Прикладом СППР на основі моделей є програмне забезпечення Oracle Hyperion Planning — система фінансового планування, яка допомагає організаціям створювати та управляти фінансовими прогнозами, бюджетами та планами. Програмне забезпечення використовує фінансові моделі та симуляції для прогнозування впливу різних фінансових рішень на фінансові показники організації. Воно надає такі функції, як моделювання сценаріїв, аналіз «що було б, якби» та прогнозування, які можуть допомогти особам, що приймають рішення, оптимізувати свої фінансові рішення та покращити бізнес-результати (*Офіційний сайт Oracle*).

4) СППР на основі знань: цей тип СППР зосереджується на використанні систем, що ґрунтуються на знаннях, для підтримки прийняття рішень. Він використовує штучний інтелект та експертні системи для надання особам, які приймають рішення, рекомендацій, заснованих на базі знань правил і фактів. СППР, на знаннях, особливо корисна для прийняття рішень у таких сферах, як охорона здоров'я, юриспруденція та фінанси (*Sharda, Delen, & Turban, 2014*).

Одним із прикладів СППР на основі знань є система медичної діагностики, яка використовує експертні знання для допомоги лікарям у встановленні точних і своєчасних діагнозів. Наприклад, лікар, може, використовувати систему, що ґрунтується на знаннях, та досвіді всіх досвідчених лікарів світу, для введення симптомів пацієнта, його історії хвороби та іншої відповідної інформації в систему, яка, у свою чергу, застосовує медичні знання і правила для генерації списку можливих діагнозів. Система, може, надавати висновки та рекомендації, які допомагають лікарю приймати обґрунтовані рішення та покращувати результати лікування пацієнтів.

Прикладом такої системи є система підтримки прийняття діагностичних рішень Isabel Healthcare, яка є веб-системою медичної діагностики, що допомагає медичним працівникам діагностувати складні захворювання. Система використовує велику базу медичних знань і правил для аналізу даних про пацієнта й генерування списку.

Система прийняття рішень — це складний процес, який включає в себе різні етапи, зокрема: ідентифікацію проблеми, збір інформації, аналіз даних, розробку альтернатив, вибір найкращого варіанту, реалізацію прийнятого рішення. Ефективність системи прийняття рішень залежить від якості інформації, аналізу та аргументації, які використовуються для прийняття рішення.

Існує декілька фреймворків для систем підтримки прийняття рішень, далі буде розглянуто модель авторів Sharifi й Retsios, (2003), які використали його у своїй роботі під назвою «Вибір місця для захоронення відходів за допомогою просторового багатокритеріального аналізу рішень». СППР у їх роботі у свою чергу, є адаптацією системи, яка була вперше згадана в роботі автора Simon, (1960) під назвою: «Нова наука про управлінські рішення». Дана модель включає чотири основні фази:

- фаза розвідки, на якій система описується і розуміється, потім визначається проблема і ставляться цілі; *(Sharifi, Retsios, 2003)*
- фаза проектування, на якій після формулювання і перевірки моделі генеруються здійсненні альтернативи формулювання та перевірки моделі; *(Sharifi, Retsios, 2003)*
- фаза вибору, на якій оцінюються альтернативи та обирається одна з них; *(Sharifi, Retsios, 2003)*
- фаза реалізації *(Sharifi, Retsios, 2003)*

Якщо зобразити всю структуру системи підтримки прийняття рішень, то можна одержати чітку послідовність дій, щодо створення такої системи приклад цих кроків показано на (рис. 1.1)

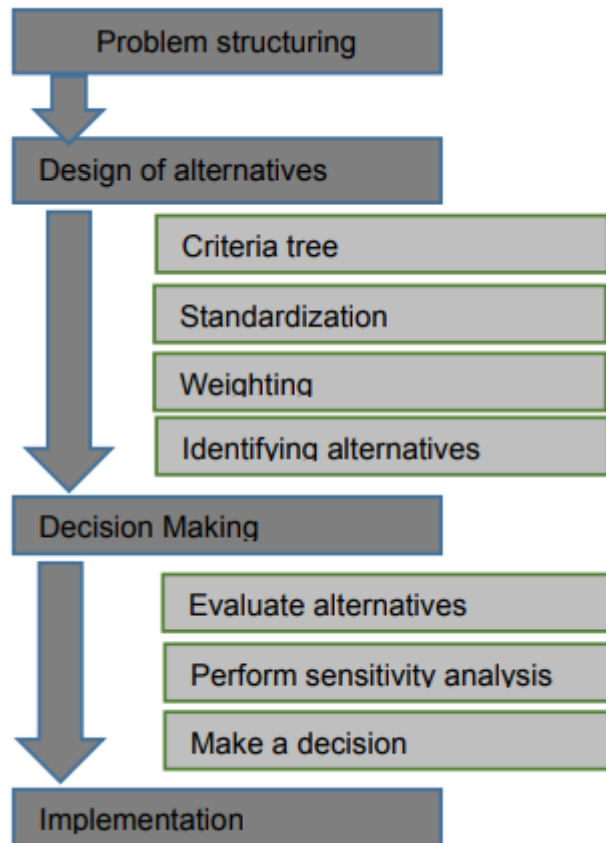


Рис.1.1 Прикладна структура СППР (Gheorghe, Armas,2015).

Відповідно до даної схеми першою сходинкою для розроблення системи прийняття рішень є ідентифікація проблеми, яка передбачає визначення питання, що потребує вирішення. Це дуже важливий крок, який задає тон усьому процесу прийняття рішень. Без чіткого розуміння проблеми неможливо прийняти обґрунтоване рішення. Автори роботи «Розумний вибір: Практичний посібник для прийняття кращих рішень». Вважають, що ідентифікація проблеми включає три основні кроки: визначення проблеми, розуміння її причин та наслідків, а також встановлення її важливості (Hammond, Keeney, and Raiffa, 1999).

Визначення проблеми вимагає чіткого і стислого опису проблеми. Для ефективного визначення можливих шляхів розв'язання проблеми важливо

розуміти її масштаби, вплив та нагальність. Розуміння причин і наслідків проблеми допомагає виробити комплексне розуміння ситуації. На цьому етапі необхідно визначити, які чинники сприяють виникненню проблеми та, які наслідки вона має. Визначення важливості проблеми допомагає розставити пріоритети в процесі прийняття рішень. Він передбачає оцінювання впливу проблеми на організацію або деяких осіб та визначення рівня її нагальності (*Hammond, Keeney, and Raiffa, 1999*).

Наступним кроком у межах цієї сходинки є збір відповідної інформації. Цей етап передбачає збирання, фактів, думок і поглядів із різних джерел, включно з експертами, зацікавленими сторонами та іншими відповідними сторонами. Зібрана інформація повинна бути точною, достовірною та релевантною до проблеми, що розглядається. У своїй роботі «Прийняття рішень: Психологічний аналіз конфлікту, вибору та зобов'язань» автори описують важливість збору інформації в процесі прийняття рішень. Вони припускають, що збір інформації, може, допомогти деяким особам і організаціям зрозуміти проблему, визначити можливі рішення і оцінити потенційні результати кожної альтернативи. Збір інформації також, може, допомогти у визначенні будь-яких потенційних ризиків або викликів, які можуть виникнути на етапі реалізації. Етап збору інформації часто є трудомістким і складним процесом, але він необхідний для того, щоби процес прийняття рішень ґрунтувався на надійних і релевантних даних. Зібрана інформація, може, надходити з різних джерел, включно з внутрішні та зовнішні (*Janis and Mann, 1977*).

Після того, як інформація зібрана, її аналізують, щоби виявити закономірності, тенденції та взаємозв'язки. Це допомагає краще зрозуміти проблему й розробити можливі рішення. Аналіз передбачає використання різних інструментів і методів, зокрема, статистичного аналізу, дерев рішень та аналізу витрат і вигод (*Tversky & Kahneman, 1974*).

Тверські та Канеман у своїй роботі від 1974 року під назвою: «Судження в умовах невизначеності: евристика та упередження». Провели впливове дослідження щодо прийняття рішень, яке показало, що люди часто приймають рішення на основі евристик, або ментальних ярликів, а не на основі раціонального аналізу. Однією з таких евристик є евристика доступності, яка передбачає прийняття рішень на основі легкості, з якою відповідні приклади спадають на думку. Це означає, що люди схильні переоцінювати ймовірність подій, які легко пригадати, наприклад, яскравих або недавніх, і недооцінювати ймовірність подій, які менше запам'ятовуються або є більш абстрактними. Дослідження також показало, що люди схильні уникати ризику, коли приймають рішення, пов'язані з вигодою, але більш схильні до ризику, коли приймають рішення, пов'язані з втратами (*Tversky & Kahneman, 1974*).

Вказана робота авторів мала значний вплив на сферу систем підтримки прийняття рішень (СППР), оскільки вона підкреслила важливість розуміння того, як люди приймають рішення і, як когнітивні упередження можуть впливати на процеси прийняття рішень. З огляду на ці знання при розробленні СППР, розробники можуть створювати більш ефективні системи, які допомагають особам, що приймають рішення, приймати поінформовані та раціональні рішення (*Tversky & Kahneman, 1974*).

Наступною сходинкою в створенні системи прийняття рішень є розроблення альтернатив. Згідно з оригінальною моделлю процесу прийняття рішень, запропонованою Simon, у роботі 1960 року, це передбачає генерування та оцінювання можливих варіантів дій. Альтернативи мусять ґрунтуватися на зібраній та проаналізованій інформації, а також бути здійсненими, практичними та ефективними в розв'язанні проблеми, що розглядається.

Для розроблення альтернатив особи, які приймають рішення, можуть використовувати різні інструменти й методи, такі, як мозковий штурм, методи номінальних груп або метод Дельфі (*Meredith & Mantel, 2012*). Зокрема, мозковий штурм, який передбачає генерування великої кількості ідей без критики чи оцінювання, а потім добір найбільш перспективних із них для подальшого аналізу. Або, наприклад, метод номінальних груп, який передбачає проведення структурованих групових сесій, під час, яких учасники генерують ідеї та визначають їхню пріоритетність, тоді, як метод Дельфі передбачає використання серії опитувальників для збору та уточнення думок експертів.

Перед останньою сходинкою в розробленні системи підтримки прийняття рішень є етап вибору моделі. Після того, як альтернативи згенеровані, особи, які приймають рішення, повинні оцінити їх за певними критеріями, які були встановлені на попередньому етапі. Це передбачає порівняння потенційних вигод і витрат кожної альтернативи, а також оцінювання їхньої реалістичності та ризиків (*Kumar & Reinartz, 2012*). Мета полягає в тому, щоб визначити альтернативу, яка найкраще відповідає цілям особи, що приймає рішення, і розв'язує проблему в найбільш ефективний і результативний спосіб.

Одним із важливих інструментів, який можна використовувати на етапі оцінювання та контролю, є матриця рішень. Матриця рішень — це інструмент для та порівняння альтернатив на основі набору критеріїв. Матриця дає змогу особам, які приймають рішення, присвоювати вагу різним критеріям і ранжувати альтернативи на основі того, наскільки добре вони відповідають цим критеріям (*Nutt, Wilson 2010*). Використовуючи матрицю рішень, особи, які приймають рішення, можуть легше оцінити ефективність обраної альтернативи та визначити сфери для вдосконалення.

Іншим важливим аспектом етапу оцінювання та контролю є постійний моніторинг і зворотний зв'язок. Це передбачає збір даних і зворотного зв'язку

від зацікавлених сторін та використання цієї інформації для коригування і вдосконалення процесу реалізації з плином часу. Постійний моніторинг і зворотний зв'язок можуть допомогти особам, які приймають рішення, виявити потенційні проблеми і виклики на ранніх стадіях і вжити коригувальних заходів до того, як вони перетворяться на серйозні проблеми (*Power, 2002*).

Останнім кроком у реалізації системи підтримки прийняття рішень є реалізація рішення. Це передбачає втілення обраної альтернативи в життя та моніторинг результатів, щоби переконатися, що заплановані результати досягнуті. Ефективна імплементація вимагає ретельного планування та координації, а також ефективної комунікації та співпраці між усіма зацікавленими сторонами, залученими до процесу прийняття рішень (*Galliers & Leidner, 2014*).

На етапі імплементації важливо відстежувати результати обраної альтернативи та вносити необхідні корективи. Це передбачає відстеження ключових показників ефективності (key performance indicators — KPI), які були обрані для вимірювання результативності в рамках другого етапу при виборі моделі, а також використання цієї інформації для оцінки ефективності рішення. Якщо результати не відповідають очікуванам, може, виникнути необхідність для того, щоби повернутися до попередніх етапів процесу прийняття рішень, наприклад, повторно проаналізувати дані або генерувати нові альтернативи та вибір нової (*Power, 2002*).

У контексті систем підтримки прийняття рішень (СППР) база даних відіграє центральну роль на етапі впровадження. Вона забезпечує централізоване сховище для зберігання та організації всіх відповідних даних та інформації, пов'язаних із процесом прийняття рішень. Вона дає змогу особам, які приймають рішення, здобувати доступ до інформації і аналізувати її в режимі реального часу, а також використовувати її для управління процесом імплементації (*Alter, 2013*). Крім того, база даних, може, бути використана для

відстеження прогресу імплементації, а також для виявлення будь-яких питань або проблем, які можуть виникнути.

У даному підрозділі були розглянуті різні типи СППР та їхня структура.

Варто підкреслити важливість добре структурованої бази даних для системи підтримки прийняття рішень у всіх галузях, зокрема, і геології. Типи СППР, які зазвичай використовуються в геології, включають просторові системи підтримки прийняття рішень, експертні системи й нейронні мережі. Ці системи покладаються на різні типи даних, включаючи геологічні мапи дані дистанційного зондування та лабораторні аналізи. Добре організована база даних має важливе значення для інтеграції інформації із цих різних джерел і для проведення всебічного аналізу з метою підтримки прийняття обґрунтованих рішень.

Важливість баз даних для таких систем у геології неможливо переоцінити. Далі буде розглянуто більш детально роль баз даних та їх важливість у підтриманні різних типів СППР та їх структур. Добре організована база даних необхідна для інтеграції даних із різних джерел, проведення всебічного аналізу та підтримки прийняття обґрунтованих рішень у галузі геології.

1.2 Огляд стану застосування СППР та баз даних в геології

Системи підтримки прийняття рішень (СППР) набувають все більшого значення в геології, допомагаючи геологам аналізувати складні масиви даних, приймати обґрунтовані рішення та оптимізувати управління ресурсами.

Одне з основних застосувань таких систем в геології - це управління мінеральними і вуглеводневими ресурсами. З їх допомогою можна аналізувати геологічні та геофізичні дані, виявляти потенційні родовища та оцінювати запаси. СППР наприклад, може, допомогти геологам скласти геологічну карту місцевості, визначити ділянки, що становлять інтерес для

подальшої розвідки, і оцінити потенційну кількість наявних мінеральних або вуглеводневих ресурсів. Також можна використовувати для оптимізації процесів буріння і видобутку, допомагаючи знизити витрати і підвищити ефективність.

Автори у своїй статті представляють систему підтримки прийняття рішень в реальному часі, яка використовує ансамблевий фільтр Калмана для апроксимації невизначеності в геологічній інтерпретації. СППР розглядає всі безпосередні альтернативи і обирає найкраще прогнозоване значення з усіх реалізацій. Система протестована на синтетичних кейсах у шаруватому геологічному середовищі і може успішно виконувати геонавігацію та посадку свердловини для різних геологічних конфігурацій об'єктів буріння. (*Alyaeu, Suter, Bratvold, Hong, Luo, & Fossum, 2019*).

СППР була протестована синтетичних прикладах з двома нафтоносними піщаними шарами, оточеними фоновими сланцями. Цільова точка буріння вибирається динамічно на основі багатоцільової функції вартості (*Alyaeu, Suter, Bratvold, Hong, Luo, & Fossum, 2019*).

В цілому СППР застосовуються будь-де так в своїй статті авторами було описано попередні кроки щодо розробки прототипу системи підтримки прийняття рішень, спрямованої на управління ресурсами підземних вод у масштабі басейну (регіон річок Тічино та Адда). Дана Система підтримки прийняття рішень є пакетом, що дозволяє оцінювати водні ресурси, визначати граничні умови, кліматичні рушійні сили та демографічний тиск, моделювати та досліджувати майбутні прогнози та порівнювати альтернативні політичні заходи (*Gorla, Simonetti, & Righetti, 2016*).

Проект був розроблений поетапно, включаючи створення бази географічних даних, аналіз географічної інформаційної системи (ГІС) (в тому числі багат шаровий аналіз) та чисельне моделювання. Дані, зібрані на базі географічних даних, були проаналізовані для розробки кількісних та якісних

тематичних карт ГІС з метою проведення багат шарового аналізу поточного та майбутнього стану і впливів, щоб надати особам, які приймають рішення, повну картину стану водної системи (*Gorla, Simonetti, & Righetti, 2016*).

Багаторівневий аналіз в роботі даних авторів спирається на специфічні показники, що базуються на певних кількісних та якісних даних: гідрогеологічні, хімічні, ізотопні, використання та небезпеки ґрунтів, кліматичні та демографічні. Кожен параметр, що належить до цих макрообластей, був класифікований за шкалою з 7 класів критеріїв, і кожному з них була присвоєна вага (*Gorla, Simonetti, & Righetti, 2016*).

Як було сказано в роботі багаторівневий аналіз є дуже корисним, сучасним та різностороннім, в даній роботі це було застосовано методом моделювання та прогнозування. А саме було створено різного роду карти, які і були цими шарами. Не завжди на основі даних можна побудувати шарувату модель такого типу. Можна створити модель подібного типу на основі інших даних, для цього потрібно застосувати багатовимірну модель баз даних.

Ще одні автори у своїй роботі описують система підтримки прийняття рішень, яка була розроблена для визначення пріоритетів у розвідці нафти і газу. Їхня система ґрунтується на знайомих концепціях теорії прийняття рішень, але призначена для використання персоналом, незнайомим з цими технічними питаннями. Система була розроблена, щоб допомогти менеджерам з розвідки у великій нафтовій компанії зробити реалістичний вибір розвідувальних проектів у своєму регіоні. Вона була інтуїтивно зрозумілою і не вимагала надмірних витрат часу та збору даних. Незбалансованість дослідження полягала в тому, що основна увага приділялася геологічній природі родовища, а інші міркування залишалися поза увагою. Геологорозвідувальний персонал регулярно оцінював обсяги очікуваних запасів на родовищах, але менеджер не міг сказати, що було скориговано, а що ні. Головна задача геологів і геофізиків на

геологорозвідувальні проекти по суті полягає в тому, щоб максимізувати цінність інформації, отриманої їхніми зусиллями. Цінність інформаційної моделі можна оцінити, використовуючи, наприклад атрибути з ігор, знайомі геологам і геофізікам (*Dyer, Lund, Larsen, Kumar and Leone, 1990*).

Реалізація підходу авторів забезпечила перше використання справжньої системи підтримки прийняття рішень на базі персонального комп'ютера в той час. А саме для допомоги в прийнятті рішень щодо розвідки нафти в цій великій нафтовій компанії. Основним принципом розробки їхньої системи була інтеграція інтуїтивних уявлень і міркувань геологорозвідувального персоналу з теоретично обґрунтованою моделлю, яка забезпечила б раціональну основу для ранжування нафтових покладів і полегшила б розподіл геологорозвідувального персоналу. Ця мета була досягнута шляхом створення наближеної лінійної моделі, яка точно відтворювала результати більш складної нелінійної моделі, отриманої на основі традиційної теорії прийняття рішень. Авторами було обрано використання регресійного аналізу, і подальші дослідження показали, що він дає хороші результати за прийнятих припущень. Важливим аспектом цієї програми була наявність персонального комп'ютера, який був відносно новою технологією на момент завершення роботи над системою на початку 1984 року. Гнучкість цього нового інструменту аналізу забезпечила зручний засіб для аналізу суджень геологорозвідувального персоналу щодо потенційних покладів нафти. Складність цього завдання унеможливила б використання цього більш формального процесу оцінки без персонального комп'ютера. СППР була використана для двох експертиз проектів з прийнятними результатами. Зі зростанням цін на нафту відновиться активне буріння відновиться активне буріння. Ця СППР може бути використана для підтримки агресивної геологорозвідувальної діяльності, коли геологорозвідувальні перфоратори

знову стануть дефіцитним ресурсом. буріння знову стане відносно дефіцитним товаром (*Dyer, Lund, Larsen Kumar and Leone, 1990*).

З оглянутих в рамках даного розділу робіт й в цілому можна зробити проміжний висновок, про те що системи підтримки прийняття рішень реалізуються лише для того, щоб передати головну суть нелінійних моделей більш простими та вже на їх основі прийняти рішення. Тобто простими словами розповісти складне. Яскравим прикладом цього є попередня робота. Також багато робіт, які пов'язані з геологією використовують дану галузь науки, як частину великого цілого, а не як частину своїх досліджень. Часто це пов'язане з оптимальним використанням ресурсів і подібними речами.

Інші автори у свою чергу основною метою свого дослідження ставили застосування просторової системи підтримки прийняття рішень для пошуку відповідного укриття в історичному центрі міста Бухарест на етапі після катастрофи, у випадку землетрусу. Їх робота є першим кроком у застосуванні просторової системи підтримки прийняття рішень в контексті сейсмічного ризику в Бухаресті. Їхній метод включав в себе наступні етапи: структурування проблеми у вигляді дерева рішень, застосування методів стандартизації та зважування критеріїв, пошук відповідних альтернатив і вибір однієї з них. альтернатив і вибір однієї з них (*GHEORGHE, ARMAS, 2015*).

Вони прийшли до висновку, що просторова система підтримки прийняття рішень інтегрує системи управління базами даних, графічне відображення, можливості табличної звітності та експертні знання осіб, які приймають рішення, для підтримки процесу прийняття рішень. Вона здатна вирішувати складні просторові проблеми. Результати показують, що кілька будівель можуть бути використані як притулок (*GHEORGHE, ARMAS, 2015*).

В даній роботі багатошаровість для прийняття рішень також була досягнута за рахунок побудов та накладання карт одна на одну. На мою думку це є досить одноманітним, хоча і ефективним. Важко сперечатися з тим, що

таких підхід важкого реалізувати у всіх сферах. Зокрема для досліджень електротомографії.

Дуже багато робіт використовують підхід поєднання геоінформаційних систем та геології для моделювання шаруватої структури геологічного середовища. Також можна використати підхід створення цифрового двійника та моделювати різного роду процеси зокрема і геологічні. Так автори роботи в якості шаруватої структури, для зображення всіх можливих шарів у видобувній індустрії, як демонстрування розвитку індустрії 4.0.

Ними пропонується багаторівнева цифрова платформа-близнюк, яка уможливить управління життєвим циклом активів, що сприятиме сталому розвитку гірничодобувної галузі, завдяки використанню еталонної архітектури моделі галузі 4.0 (з англ. RAMI - Reference Architectural Model Industrie), управлінню активами протягом усього їхнього життєвого циклу та створенню спільної екосистеми, що включає чотири сервіси (*Bazi, Mabrouki, Laayati, Ouhabi, Hadraoui, Hammouch та інші, 2023*)

У їх роботі представлено еталонний багаторівневий підхід до рамкової архітектури цифрового двійника (DT- digital twins) в Індустрії 4.0 для гірничодобувної промисловості, а також обговорюється рівень інтеграції технології DT та її потенціал для забезпечення управління життєвим циклом вартості. (*Bazi, Mabrouki, Laayati, Ouhabi, Hadraoui, Hammouch та інші, 2023*)

Їх дослідження мало на меті створити цифрову структуру-двійника для управління життєвим циклом активів у гірничодобувній промисловості на основі результатів попереднього огляду літератури та фокус-груп. Система була розроблена таким чином, щоб бути дуже гнучкою та адаптованою до конкретних умов і викликів гірничодобувної галузі. (*Bazi, Mabrouki, Laayati, Ouhabi, Hadraoui, Hammouch та інші, 2023*)

Еталонна модель архітектури для Індустрії 4.0 (RAMI 4.0) може допомогти спростити складні завдання до пакетів, які добре узгоджуються з ключовими

характеристиками Індустрії 4.0. Така модель складається з тривимірної системи координат. Де кожна розмірність це кожний параметр даної моделі, а не географічні координати. Ця модель спрощує складні взаємозв'язки до менших, більш керованих кластерів. (*Bazi, Mabrouki, Laayati, Ouhabi, Hadraoui, Hammouch та інші, 2023*).

На мою думку даних підхід до моделювання можна застосувати не тільки для бізнес сторони геології, а й до самого геологічного середовища. Що і було зроблено авторами в наступній роботі. Вони вважають, що геометрична точка зору геологічних моделей може забезпечити зв'язок між геологічними і геофізичними спостереженнями з додатковою геологічною і фізичною інформацією та знаннями, оскільки багато відповідних аспектів поділяють концепцію дискретних інтерфейсів у надрах як об'єднуючий елемент. Як приклад для дослідження вони взяли область складчастого карбонатного шару в Піренеях (Арагон, Іспанія), який пояснюється осадовими і тектонічними процесами (*Wellmann, Caumon, 2018*).

Людство, та людина як особа не живуть вічно й через не може відобразити всю складну історію кількісно. Але за допомогою різних досліджень ми можемо вловити суть геологічних подій у тривимірному просторі, таких як зміни в системі осадконакопичення або деформаційні події, такі як розломи або в'язкі деформації. Фундаментальною метою геологічного моделювання є отримання методів, які дозволяють обґрунтовано описувати геометричні та топологічні особливості в повному тривимірному просторі. Побудова геологічної моделі обмежується здатністю обраних методів відображати рівень геологічного реалізму. Оскільки інформація та знання обмежені, опис може бути локально неточним. Невеликий розлом, видимий у відслоненні, відсутній у моделі (*Wellmann, Caumon, 2018*).

Дані автори виділяють в своїй роботі дуже багато різних підходів, щодо геологічного моделювання. Оскільки саме виділення методик та підходів було

їх головною метою. Варто просто згадати декілька з них. Зокрема, це підхід, заснований на даних. Вони вважають, що усі підходи на основі геофізичних даних повинні починатися з геофізичних вимірювань, які можуть бути виконані з певною метою або як загальні регіональні вимірювання. Такі дослідження можна використовувати для отримання перших уявлень про надра, але основний внесок зазвичай отримують після спеціальної обробки даних. (Wellmann, Caumon, 2018).

В цілому я вважаю, що з ними важко не погодитися. Також при обробці сейсмічних даних геологічні знання також можуть входити в робочий процес обробки на цьому етапі, наприклад, у вигляді меж для швидкісних моделей міграції при обробці сейсмічних даних або при аналізі даних про потенційні родовища, які також можуть безпосередньо враховувати геологічні моделі (Wellmann, Caumon, 2018).

Також ними було виділено підхід на основі геомodelей. Автори прийшли до думки, що для розуміння дослідження надр необхідна інша точка зору, саме тому використовуються геологічні карти, а також тривимірні геологічні моделі. З цієї точки зору, геологічні спостереження і вимірювання є первинними даними, але їх ніколи не можна відокремлювати від геологічних знань (Wellmann, Caumon, 2018).

Автори зазначають, що вхідні дані для геомodelювання включають дебіти свердловин, вимірювання перехідного тиску/висоти та складу флюїдів, а також позначені точки, що вказують на найбільш ймовірне місцезнаходження та/або орієнтацію певної геологічної поверхні (Wellmann, Caumon, 2018).

В рамках даної роботи вхідними даними слугували дані вимірювання 2Д профілів методом електротомографії та дані про прилад, а також оператора. Решта даних для моделювання відсутні. Оскільки даний метод має порівняно не велике проникнення тому важко сказати, які саме породи він охоплював,

та багато інших факторів є не відомими, але для моделювання даної системи це не грає великої ролі.

Також вище згаданої роботи зазначають, що у всіх геологічних моделях присутні не визначеності. Вони завжди були невід'ємним аспектом геологічних досліджень і відображені в найдавнішій формі геологічних моделей - геологічних картах. Зокрема ними було виділено наступні не визначеності: не визначеність в контексті побудови моделі; не визначеність даних та вимірювань; не визначеності пов'язані з різними типами даних. Відповідно з ними потрібно робити розподіл ймовірностей для різних джерел невизначеностей, а також залучати експертів. Також в 3Д моделюванні можуть бути присутніми топологічні невизначеності та об'єднання даних (m , Φ , k , β та τ) (Wellmann, Caumon, 2018).

Також шарувату геологічну модель можна зобразити використовуючи бази даних. Наприклад наступні автори в своїй статті ставлять основною метою показати, як можна управляти геологічними, інженерно-геологічними та геотехнічними знаннями в добре організованій базі даних, яка здатна розвивати різні кореляції і представляти корисну інформацію як на стадії проектування, так і на стадії будівництва важливих проектів цивільного захисту, муніципалітету і технічних компаній (Kokkala, Marinos, 2022).

Своїм дослідженням вони демонструють виконання такої роботи в місті Салоніки на півночі Греції, де велика кількість геологічних і геотехнічних даних для більш ніж 600 свердловин, в основному на основі свердловинної геологічної інформації, лабораторних випробувань, геотехнічної характеристики і польових випробувань на місці, зібраних із заповнювача, четвертинних і неогенових відкладів, зберігається в базі даних з прив'язкою до місцевості. Створена ними база даних призначена для управління та обробки наборів геоінженерних даних для різних геотехнічних проектів, а також для захисту від стихійних лих (Kokkala, Marinos, 2022).

У своїй роботі дані автори зібрали великі масиви даних з різних джерел, обробили та використали у великій міській місцевості для розвитку знань і вдосконалення геотехнічних послуг розумного міста. Велика кількість наборів даних були взяті з більш ніж 600 свердловин з геологічною інформацією, в основному даними натурних і лабораторних випробувань, зокрема, таких, як: стандартний тест на проникнення, міцність на необмежений стиск, межі Аттерберга, індекс консистенції і природна вологість, автори розробили систему їх зберігання в інженерно-геологічній базі даних. Їхні набори даних аналізуються і корелюються для ретельного визначення складного характеру земної поверхні міста Салоніки (*Kokkala, Marinos, 2022*).

Метою такого підходу автори вважають є, окрім самої інтеграції наборів даних з різних джерел, є надання інструменту для легкого і безпосереднього аналізу, кореляції, а також створення таблиць і діаграмах для зацікавлених сторін є відображення даних на картах (*Kokkala, Marinos, 2022*).

Також за допомогою створення такої бази даних авторам запропоновано конкретні діапазони значень для кількох фізико-механічних властивостей, що забезпечує геологічні та геотехнічні знання щодо безпечного розширення міст, придатних для будівництва в умовах досліджуваної території (*Kokkala, Marinos, 2022*).

З даної роботи можна зробити висновок, що не тільки інтеграція геоінформаційних систем разом з геологією дає нову інформацію, а ще й створення баз даних, на основі, яких можна отримати нові знання про ту чи іншу область.

Таким чином на етапі вибору альтернати в даній роботі було зроблено вибір на створенні бази даних з декількох причин. По-перше простота реалізації, підтримки та використання в подальшому оскільки це не потребує багато знань з різних областей. Використання досить просто достатньо вміти складати запити мовою SQL. По-друге мала кількість даних. Звісно можна було

зробити інтегрування ГІС разом з вхідними даними, але для цього потрібні різного роду карти місцевості на якій проводили вимірювання.

Але серед баз даних є багато різновидів, зокрема, як відмінності в їх схема де є реляційні та багатовимірні чи з використанням мови SQL або без неї. Також серед всіх цих видів є різні системи керування база даних, як обрати серед них правильну важко сказати, оскільки потрібно орієнтуватися у вхідних даних та їх особливостях.

Автори роботи під назвою «Універсальна та багатовимірна модель для аналітичних даних про геологічних зразків». Відповідно до назви ставлять собі на меті розробку нової модель геоданітичних даних, використовуючи різного роду геологічні дані. Незважаючи на свої обмеження, на сьогоднішній день реляційна модель даних є найбільш поширеною моделлю для геоданітичних даних (He, Tian, Wang, Yao, Chen, 2019).

Реляційна модель даних будує зв'язки між кожною групою даних у базі даних. Це в свою чергу означає, що більша кількість категорій даних неминуче призводить до набагато більшої кількості зв'язків між даними, збільшуючи вимоги до сховища і час, необхідний для запитів до бази даних для запитів до бази даних. У порівнянні з такими традиційними моделями даних, багатовимірні моделі даних (Multidimensional data model - MDM), які широко використовуються під час розвитку науки про великі дані та інтелектуального аналізу даних, є єдиним предметно-орієнтованим джерелом для аналізу даних на основі різних вимірів. Підходи багатовимірного моделювання мають спільні характеристики з швидким аналізом спільної багатовимірної інформації (Fast Analysis of Shared Multidimensional Information, FASMI). Зокрема, MDM пропонує перевагу відносно простого і зрозумілого дизайну бази даних, яка, тим не менш, підтримує потужний аналіз і є відносно зрозумілою для кінцевих користувачів. Як система моделювання, багатовимірна модель даних має концептуальну і логічну фазу проектування,

що складається з таблиці фактів і декількох таблиць вимірів. Факти включають числові та адитивні характеристики даних, які можуть бути накопичені за кількома вимірами. Часто дослідники зацікавлені в аналізі геоаналітичних даних вимірювань з різних перспектив метаданих. Отже, на думку дослідників підхід MDM ідеально підходить для розробки геоаналітичних моделей даних. Також в кінці своєї роботи автори приводять порівняльні графіки, які чітко стверджують, що багатовимірна база даних, є кращою з технічної точки зору, зокрема в швидкості обробки запитів, а також в простоті їх складання (*He, Tian, Wang, Yao, Chen, 2019*).

Є роботи, які на прикладі видобувної сфери доводять, те що мультимімерні бази даних в даній області є чудовим варіант, але у всього є свої особливості і така схема не є виключення. Зокрема в роботі «Аналіз джерел даних та вимог до них для Багатовимірне моделювання баз даних – австралійський сценарій для видобувної промисловості Австралії» вказують, що є наступні вимоги, як для самих даних так і для їх моделі, а саме: аналіз вимог до даних для багатовимірного підходу, також повинна бути симетрія вимірів і мір. Дана особливість полягає в тому, що виміри можна розглядати як міри і навпаки, і може існувати більше одного шляху для агрегування даних у кожному вимірі. Модель даних повинна підтримувати правильне агрегування, наприклад, підрахунок одного і того ж відкриття нафти або газу один раз на групу, навіть якщо відкриття нафти або газу зроблено в декількох басейнах. (*Nimmagadda, Rudra, 2004*).

Автори вважаю, що багатовимірне моделювання - це метод проектування баз даних для розробки сховищ даних і пов'язаних з ними додатків, в якому концептуальні моделі переводяться в розмірні моделі, класифікуючи розмірності, розробляючи високорівневі схеми типу «зірки», створюючи дрібні таблиці розмірностей і фактів. У багатовимірному поданні кожна комірка містить дані, що відносяться до перетину всіх значень вимірів.

Еквівалентне представлення також називається зірковою схемою і складається з розмірних таблиць з центрально розташованою таблицею фактів. А також вони підтверджують отримані висновки, які були згадані в попередній роботі, про те, що багатовимірне моделювання оптимізує продуктивність запитів завдяки використанню зоряної схеми та різних рівнів узагальнення даних (*Nimmagadda, Rudra, 2004*).

З оглянутих робіт можна зробити висновок, про те що досить потужним методом моделювання геологічного середовища, як шаруватої структури є багатовимірна модель бази даних. Мною також було оглянуто багато інших праць. І здійснено висновки про те, що такі схеми в геології дуже рідко використовуються. Застосовуються лише для моделювання сейсміки, як частина системи підтримки прийняття рішень, зокрема, як було згадано для пошуку оптимального сховища, чи моделювання області зсувів. Також є деяка кількість робіт де таку схему застосовують для зберігання сейсмічних даних, які були отримані самими дослідниками, але такі схеми часто обмежуються мало кількістю розмірностей.

Якщо ж окремо говорити про те, що використовується для зберігання даних отриманих в результаті електротомографії Землі, це просто жах. Складається враження ніби дослідники взагалі не знають, що таке багатовимірні моделі баз даних, та той факт, що їх дуже легко створити, використовувати, та обслуговувати. Загалом, майже, з впевненістю можна сказати, що такі бази даних для таких даних чомусь не застосовуються, для цих даних. Зазвичай дослідники створюють стільки розмірностей, стільки мають самих даних.

Але деякі багатовимірні моделі все ж таки вченими було виконано зокрема 3Д модель на основі даних електротомографії. Але це було досягнуто зовсім іншими методами, більш традиційними для математичного моделювання. Зокрема, в своїй роботі під назвою: «Багатовимірне

дослідження електричного опору для виявлення корінних порід на Рівнина Рієті (Центральна Італія)». Використали комбінування методів одномірного, двомірного і тривимірного електричного опору для оцінки глибини залягання корінних порід, які було представлено на прикладі конкретного дослідження, проведеного в Центральній Італії. В рамках даного дослідження авторами було виконано шість вертикальних електричних зондувань вздовж поперечного перерізу басейну з метою попередньої 1D реконструкції корінних порід та алювіальних відкладів, що перекривають їх. В їх роботі інвертовані моделі електричного опору показують глибину залягання корінних порід відповідно до даних свердловин і складну підповерхневу шаруватість розкритих відкладів, що підлягають подальшому дослідженню за допомогою 2D і 3D електричної томографії електричного опору (*Cardarelli, Donno, 2017*).

Завдяки томографічній інверсії 3D-даних в рамках своєї роботи вони змогли реконструювати об'ємне зображення карбонатного утворення на досліджуваній ділянці. Томографічні моделі були перевірені шляхом інверсії синтетичного набору даних з метою отримання остаточної моделі, синтетична модель якої є порівнянною з польовою (*Cardarelli, Donno, 2017*).

Все ж таки деякі дослідники використовують метод електротомографії для багатовимірного моделювання, але не в рамках СППР, а в рамках математичної моделі, яку було візуалізовано. Так автори роботи «Моніторинг ґрунтів за допомогою 4D електрорезистивної томографії динаміка вологості в експлуатованому залізничному насипі». Зокрема в них четвертою розмірністю виступав час, оскільки це є моніторинг тобто вимірювання проводилися з певною частотою (*Chambers, Gunn, Wilkinson, Meldrum, Haslam, Holyoake, 2014*).

В своєму дослідженні дані автори використовували 2D ERT для моніторингу зміни розподілу вологи в межах насипу і був тісно пов'язаний із

зонами з поганою геометрією залізничної колії, оскільки вона була побудована ще в 1895 році і може бути аварійною, через підвищення вологості. У цьому дослідженні вони використовували 2D і 3D ERT, а також ручний повторний і повністю автоматизований збір даних для дослідження сезонної динаміки вологості на ділянці залізничної системи. Всі дані про питомий електричний опір були зібрані лінійно з використанням дипольно-дипольної сітки, а порівняння прямих і зворотних вимірювань забезпечило надійні та кількісні критерії редагування даних (*Chambers, Gunn, Wilkinson, Meldrum, Haslam, Holyoake, 2014*).

Дане дослідження авторів є актуальним оскільки динаміка внутрішньої вологості старого залізничного земляного насипу була досліджена за допомогою 2D і 3D моніторингу питомого опору. Візуалізація даних у вигляді 2D перерізів, 3D томограм та графіків часових рядів для різних зон насипу дозволила розробити сезонні фронти зволоження в межах дамби, які можна відслідковувати з високою просторовою роздільною здатністю, а також відповідний розподіл вологості у бокових частинах, на гребені та підшві дамби, а також оцінити розподіл вологості у підшві дамби (*Chambers, Gunn, Wilkinson, Meldrum, Haslam, Holyoake, 2014*).

Авторами отримання даних 2D і 3D ERT вони були інвертовані з використанням регуляризованого алгоритму оптимізації за методом найменших квадратів, і була досягнута хороша збіжність між спостереженими і модельними даними як для 2D, так і для 3D моделей (*Chambers, Gunn, Wilkinson, Meldrum, Haslam, Holyoake, 2014*).

Переваги використання СППР в геології численні. Такі системи підвищують точність і ефективність розвідки та управління ресурсами, знижуючи витрати і збільшуючи прибуток. Вони також підвищують безпеку і знижують ризик екологічної шкоди або небезпеки. Надаючи особам, які приймають рішення, доступ до відповідної інформації та інструментів для

аналізу, вони допомагають приймати обґрунтовані рішення, які базуються на даних, а не на інтуїції чи здогадках.

Однак, використання СППР в геології пов'язане з певними труднощами. Однією з головних проблем є доступність та якість даних. Геологічні дані можуть бути складними і важкодоступними для отримання, і в них можуть бути прогалини або невідповідності, які можуть вплинути на їх точність.

Центральним елементом таких систем є база даних. До якої в свою чергу ставиться багато критеріїв. Також СППР повинна відображати шарувату структуру Землі, це можна зробити багатьма методами, які були описані в межах даного розділу. Найбільш оптимальним на мою думку є створення багатовимірної бази даних.

2. Багатовимірні бази даних

Багатовимірна база даних — це особливий тип бази даних, оптимізований для зберігання інформації та використання в OLAP (онлайн-аналітичної обробки). Багатовимірна база даних являє собою структуровану комбінацію показників із різних джерел, які одночасно працюють між БД і пропонують мережі, ієрархії, масиви та інші методи форматування інформації. У ній дані представлені користувачам у вигляді багатовимірних масивів, а кожне поодиноке значення міститься в комірці, до якої можна отримати доступ за допомогою декількох індексів (*Rouse, 2017*).

Технологія багатовимірних баз даних є ключовим фактором в інтерактивному аналізі великих обсягів показників для прийняття рішень. На відміну від інших технологій, зокрема, реляційні БД ці розглядають дані, як багатовимірні куби, які, особливо, добре підходять для аналізу даних. Багатовимірні моделі класифікують дані або, як факти з відповідними числовими показниками, або, як текстові виміри, що характеризують факти. Запити об'єднують значення вимірювань у діапазоні значень вимірів, щоб одержати такі результати, як загальний обсяг продажів певного продукту за місяць (*Pedersen, Jensen, 2001*).

Технологія багатовимірних БД застосовується до розподіленої інформації і до нових типів даних, які сучасні технології часто не можуть адекватно проаналізувати. Усе частіше вона буде застосовуватиметься там, де результати аналізу передаються безпосередньо в інші системи, усуваючи в такий спосіб людину із циклу обробки даних. У поєднанні з необхідністю безперервного оновлення, цей контекст висуває жорсткі вимоги до продуктивності, які не задовольняються сучасними технологіями (*Pedersen, Jensen, 2001*).

Тепер варто розглянути історію виникнення оскільки, якщо, зрозуміти для, чого вони із самого початку були призначені можна зрозуміти, їх можливі сфери застосування, а також їх недоліки та переваги.

Багатовимірні бази даних беруть свій початок не з технології БД, а з багатовимірної матричної алгебри, яка використовувалася для ручного аналізу даних із кінця дев'ятнадцятого століття (*Pedersen, Jensen, 2001*).

Наприкінці 1960-х років дві компанії зі США IRI Software та Comshare незалежно одна від одної почали розробляти те, що згодом стало багатовимірними системами БД. IRI Express, популярний інструмент для маркетингового аналізу наприкінці 1970-х та на початку 1980-х років, став провідним інструментом аналітичної обробки даних в Інтернеті, а згодом був придбаний компанією Oracle. Одночасно система Comshare розвинулася в System W, яка активно використовувалася для фінансового планування, аналізу та звітності впродовж 1980-х років (*Pedersen, Jensen, 2001*).

У 1991 році була створена компанія Arbor Software, нині Hyperion Solutions, з конкретною метою з конкретною метою створення багатокористувацької, багатовимірної БД багатокористувацького, багатовимірного сервера баз даних, результатом чого стала система Essbase. Пізніше Arbor ліцензувала базову версію Essbase компанії IBM для інтеграції в DB2 (*Pedersen, Jensen, 2001*).

У 1998 році автори Salley, та Codd, ввели термін OLAP в сучасному його розумінні. Ще однією значною подією на початку 1990-х років стала поява великих сховищ даних, які, як правило, базуються на реляційних схемах «зірка» або «сніжинка» — підході, що використовує технологію реляційних БД підхід, який використовує технологію реляційних баз для реалізації багатовимірних. (*Salley, Codd, 1998*).

У 1998 році Microsoft випустила свій MS OLAP Server, першу багатовимірну систему, орієнтовану на масовий ринок, і тепер багатовимірні системи стають

товарними продуктами, що поставляються без додаткових витрат разом із провідними системами реляційних баз даних (Pedersen, Jensen, 2001).

Необхідно взагалі розглянути структури таких баз, що є головним елементом, що є необхідним побічними структурами.

Головною одиницею будь-якої БД є таблиця і багатовимірна схема не є виключенням. Приклад двох вимірної таблиці з уявними даними показано нижче в (табл. 2.1).

Таблиця 2.1 – Приклад електронної таблиці товарів.

Товари	Кількість покупок по містах		
	Північна Європа	Центральна Європа	Південна Європа
Туристичне спорядження	583	825	935
Аksesуари	629	425	534
Обладнання для гольфу	353	562	496
Альпіністське спорядження	564	826	356

Електронна таблиця, подібна до тієї, що наведена в Таблиці 2.1, є корисним інструментом для розгляду даних, наприклад, про продажі, які показано в цій таблиці. Зокрема, аналізу таких показників, як проданий товар, кількість покупок та місто продажу. Зведена таблиця — це двовимірна електронна таблиця з відповідними проміжними та загальними підсумками, яка підтримує перегляд більш складних даних, вкладаючи кілька вимірів на осі x або y та їх відображення на декількох сторінках. Зведені таблиці зазвичай підтримують інтерактивний вибір підмножин та зміну рівня деталізації даних, що відображаються. Електронні таблиці є недостатнім інструментом для управління та зберігання багатовимірних даних, оскільки вони занадто тісно прив'язують зберігання інформації надто тісно пов'язане з їхнім представленням — вони не відокремлюють структурну інформацію від бажаного уявлення про інформацію. Отже, додавання третього чи іншого виміру, наприклад, час, або групування даних за типами продуктів вищого

рівня типів продуктів вимагає значно складніших взаємозв'язків. Очевидним рішенням є використання окремої електронної таблиці для кожного виміру, але це буде працювати, лише в обмеженою мірою, оскільки аналіз додаткових значень додаткового виміру швидко стає громіздким (*Pedersen, Jensen, 2001*).

Взагалі існує кілька можливостей візуалізувати багатовимірну базу даних, один із них це куб, але він обмежений трьома, або максимум чотирьома вимірами, а якщо їх більше тоді їх можна продемонструвати у вигляді таблиць та зв'язків між ними. Тоді таке представлення буде мало, чим відрізнитися від реляційної бази, але це тільки візуально структурно та технологічно це будуть дуже різні бази даних.

Найчастіше ж представляють дані, як йшлося вище у вигляді кубів, які узагальнюють електронні таблиці до будь-якої кількості вимірів. Крім того, куби підтримують ієрархії вимірів і формул, не дублюючи їх визначення. Сукупність пов'язаних між собою кубів становить багатовимірну БД або сховище даних (data warehouse). Оскільки виміри в кубі є першокласними, вбудованими концепціями з відповідними доменами, куби можуть легко керувати додаванням нових значень вимірів. Хоча термін передбачає три виміри, теоретично куб, може, мати будь-яку кількість вимірів; насправді здебільшого у реальних кубах створюють від чотирьох до дванадцяти вимірів. Комбінації значень вимірів визначають комірки куба (*Pedersen, Jensen, 2001*).

На рисунку 2.1 показано куб, що відображає продажі для різних частин Європи з таблиці 2.1 з додатковим виміром часу, а саме розподілом по роках. У відповідних клітинках зберігається кількість продажів. У прикладі є факт — непорожня клітинка, яка містить ряд пов'язаних числових показників — для кожної комбінації часу, продукту та міста, де було здійснено принаймні один продаж. У комірках зберігаються числові значення, пов'язані з фактом — у цьому випадку кількість продажів являє собою єдину міру обрахунків. Зазвичай куб підтримує одночасне відображення, лише двох або трьох

вимірів, оскільки візуально важко уявити більше розмірностей, але він, може, показувати до чотирьох вимірів низької кардинальності, вкладаючи один вимір в інший на осях. Отже, розмірність куба зменшується під час запиту шляхом проектування його до 2D або 3D через агрегування значень вимірів у спроектованих вимірах, що призводить до отримання значень вимірів вищого рівня для бажаного перегляду даних. Наприклад, щоби переглянути продажі за певною частиною Європи й часом, потрібно агрегувати по всьому виміру товарів для кожної комбінації місця і часу. Отже, на Рисунку 2.1 додавання 629, 425 та 534 дає загальний обсяг продажів аксесуарів у всій Європі (Pedersen, Jensen, 2001).

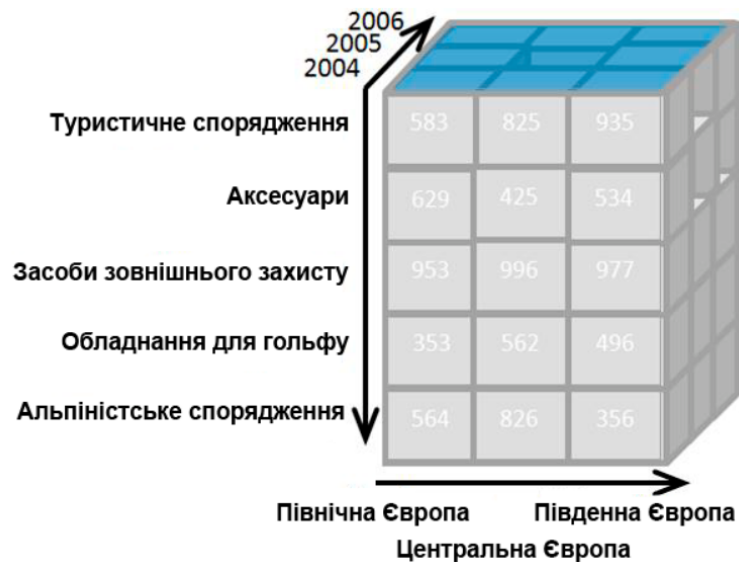


Рис. 2.1 Зображення куба, як приклад візуалізації багатовимірної бази даних (Вікіпедія ОЛАП куб, 2010)

Виміри є важливою і відмінною концепцією багатовимірних баз даних. Важливою метою багатовимірного моделювання є використання вимірів для забезпечення якомога більшого контексту для фактів. На відміну від реляційних баз даних, контрольована надмірність зазвичай вважається доречною в багатовимірних базах даних, якщо вона підвищує інформаційну цінність даних. Оскільки дані багатовимірного куба часто отримуються з інших джерел - наприклад, з транзакційної реляційної системи - а не "народжуються" в багатовимірному кубі, проблемами надмірності,

пов'язаними з оновленнями, можна легше керувати. Зазвичай немає надмірності у фактах, лише у вимірах. Виміри використовуються для відбору та агрегування даних на потрібному рівні деталізації. Вимір організовується в ієрархію, що складається з численних рівнів, кожен з яких представляє рівень деталізації, необхідний для бажаного аналізу. Кожен екземпляр виміру або значення виміру належить до певного рівня (*Pedersen, Jensen, 2001*).

Для багатовимірних моделей іноді корисно визначити кілька ієрархій для одного виміру - наприклад, модель може визначати час як фінансовий рік і як календарний рік. Кілька ієрархій мають один або кілька спільних найнижчих рівнів - наприклад, день і місяць - і модель групує їх на кілька вищих рівнів - фінансовий квартал і календарний квартал - щоб забезпечити легке посилання на кілька способів групування. Щоб уникнути дублювання визначень, метадані куба або багатовимірної бази даних визначають ієрархію вимірів. Приклад такої ієрархії показано на рис. 2.2 (*Pedersen, Jensen, 2001*).

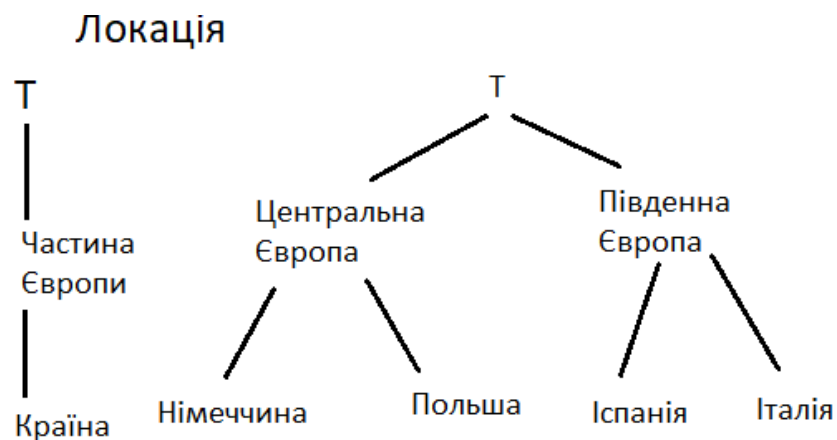


Рис. 2.2 Приклад ієрархії в багатовимірній базі даних.

Варто зазначити, що така ієрархія, може, і не застосовуватися в багатовимірних БД. Вона використовується, лише за потреби, а коли використовується дана база переходить від типу «зірка» до «сніжинки».

На Рисунку 2.2 показано схему та приклади вибіркового виміру місцезнаходження для даних про продажі в Таблиці 2.1. З трьох рівнів виміру місцезнаходження найнижчим є рівень країни, хоча, зрозуміло, що можуть бути додані ще кілька рівнів, зокрема, міста, та райони міст, або, навіть, певний

магазин. Рівень T — таблиця представляє весь вимір. У деяких багатовимірних моделях рівень має ряд пов'язаних властивостей, які містять просту, не ієрархічну інформацію. Наприклад, розмір упаковки, може, бути властивістю рівня у вимірі продукту. Вимір розміру упаковки також, може, містити цю інформацію. Використання властивості рівня не збільшує розмірність куба. На відміну від лінійних просторів, що використовуються в матричній алгебрі, багатовимірні моделі зазвичай не включають метрики порядку або відстані для значень розмірності. Єдине впорядкування полягає в тому, що значення вищого рівня містять значення нижчого рівня. Однак для деяких вимірів, таких, як час, впорядкування значень виміру використовується для обчислення кумулятивної інформації, наприклад, загальний обсяг продажів на сьогодні. Здебільшого моделі вимагають ієрархії вимірів для формування збалансованих дерев — ієрархія повинна мати однакову висоту скрізь, і кожне значення, що не є вершиною, має точно одного батька (*Pedersen, Jensen, 2001*).

Головною таблицею в такій схемі бази даних є таблиця фактів. Дана таблиця містить ті дані, які безпосередньо вимірюються у випадку з продажами, це кількість проданого товару, решта інформації надходить з інших таблиць. Тобто фактично дана таблиця, може, містити менше унікальної інформації, ніж таблиці розмірностей, але їх поєднання дає велику ефективність для аналізу даних (*Pedersen, Jensen, 2001*).

Факти представляють предмет — цікаву закономірність або подію на підприємстві, яку необхідно проаналізувати, щоб зрозуміти його поведінку. Здебільшого в багатовимірних моделях даних факти неявно визначаються комбінацією значень вимірів; факт існує, лише, тоді, коли існує непорожня комірка для певної комбінації значень. Однак деякі моделі розглядають факти, як першокласні об'єкти з окремою ідентичністю. Більшість багатовимірних моделей також вимагають зіставлення кожного факту з одним

значенням на найнижчому рівні в кожному вимірі, але деякі моделі пом'якшують цю вимогу до зіставлення (*Pedersen, Jensen, 2001*).

Кожен факт має певну деталізацію, яка визначається рівнями, з яких береться його комбінація значень виміру — наприклад, деталізацією факту куба на рисунку 2.1 є рік за продуктом за частиною Європи. Деталізація, що складається зі значень вимірів вищого або нижчого рівня, ніж задана деталізація, наприклад, рік за типом продукту, за містом і день за продуктом за містом, є грубішою або дрібнішою відповідно. Сховища даних зазвичай включають три типи фактів. Тобто одна база, може, містити декілька таблиць із фактами зокрема, такі як (*Pedersen, Jensen, 2001*):

— Події принаймні з найдрібнішою деталізацією, зазвичай моделюють події реального світу, де один факт представляє один і той самий екземпляр базового явища. Прикладами можуть бути продажі, кліки на веб-сторінці або переміщення товарів на склад і зі складу (*Pedersen, Jensen, 2001*).

— Знімки моделюють стан об'єкта в певний момент часу, наприклад, рівень запасів у магазині чи на складі або кількість користувачів веб-сайту. Один і той самий екземпляр реального явища — наприклад, конкретна банка квасолі на полиці — може, зустрічатися в декількох фактах у різні моменти часу (*Pedersen, Jensen, 2001*).

— Кумулятивні знімки обробляють інформацію про активність до певного моменту. Наприклад, загальний обсяг продажів до поточного місяця цього року, включно можна легко порівняти з показником за відповідний місяць минулого року (*Pedersen, Jensen, 2001*).

Оскільки вони підтримують взаємодоповнюючі класи аналізу, певне сховище даних часто містить усі три типи фактів. Дійсно, ті самі базові дані — наприклад, рух товарів на складі — можуть бути включені в три різні типи кубів: рух на складі, запаси й рух із початку року до сьогодні (*Pedersen, Jensen, 2001*).

У багатовимірній БД показники зазвичай представляють властивості факту, які користувач хоче оптимізувати. Потім міри набувають різних значень для різних комбінацій вимірів. Властивість і формула вибираються, отже, щоб забезпечити значуще значення для всіх комбінацій рівнів агрегації. Оскільки метадані визначають формулу, дані не реплікуються, як у електронній таблиці. Більшість багатовимірних моделей даних мають міри, але деякі з них покладаються на використання значень вимірів для обчислень за рахунок зручності для користувача. (*Pedersen, Jensen, 2001*).

Багатовимірна база даних відповідно до своєї структури піддається певним типам запитів:

1) Запити на розрізання роблять вибірку для зменшення куба. Наприклад, ми можемо розрізати куб на рисунку 2.1, розглядаючи, лише ті клітинки, які стосуються аксесуарів, а потім ще більше зменшити цей зріз, розглядаючи, лише клітинки за 2005 рік. Вибір одного значення розмірності зменшує розмірність куба, але можливі й більш загальні варіанти вибору (*Pedersen, Jensen, 2001*).

2) Запити на згортання та розгортання — це зворотні операції, які використовують ієрархії розмірностей та міри для виконання агрегації. Згортання до верхнього значення відповідає пропуску виміру. Наприклад, згортання від країни до регіону на Рисунку 2.2 об'єднує значення для Іспанії та Італії в одне значення — Південна Європа (*Pedersen, Jensen, 2001*).

3) Запити з перехресним запитом об'єднують куби, які мають один або кілька вимірів. З погляду реляційної алгебри, ця операція виконує об'єднання (*Pedersen, Jensen, 2001*).

4) Запити на ранжування або верхні n /нижні n кількість запитів можуть повернути, лише ті клітинки, які знаходяться вгорі або внизу вказаного порядку, наприклад, 10 найпопулярніших товарів в Італії у 2005 році (*Pedersen, Jensen, 2001*).

5) «Обертання» куба дозволяє користувачам побачити дані, згруповані за іншими параметрами (*Pedersen, Jensen, 2001*).

Дуже важливо також розглянути типи багатовимірних баз даних, оскільки вони можуть відрізнятися своїми особливостями, недоліками та перевагами, що, у свою чергу, робить важливим вибір правильного типу бази даних відповідно до конкретних задач. Реалізації багатовимірних баз даних мають дві основні форми: Це MOLAP та ROLAP, а також гібридну з двох основних. Розглянемо їх більш детально.

2.1 Багатовимірна аналітична обробка, як вид багатовимірної бази даних

Багатовимірна аналітична обробка в режимі онлайн зберігає дані на дисках у спеціалізованих багатовимірних структурах. (MOLAP скорочено з англійської мови — Multidimensional online analytical processing) MOLAP системи, як правило, включають положення для обробки розріджених масивів і застосовують розширене індексування та хешування для пошуку даних при виконанні запитів (*Pedersen, Jensen, 2001*).

Типова MOLAP-система містить три концептуальні компоненти: сервер БД, MOLAP-сервер і інструмент презентації, який приймає запити аналітиків і видає результати. Вони працюють разом наступним чином (*Dantoni, 2022*) :

1) У традиційній трирівневій архітектурі MOLAP сервер БД є першим рівнем. Він слугує сховищем для даних, які витягуються з різних джерел, а потім очищаються та інтегруються (*Dantoni, 2022*).

2) Середній рівень MOLAP-системи — це MOLAP-сервер, який отримує інформацію зі сховища (зазвичай через SQL) і завантажує їх у багатовимірні куби даних, оптимізовані для запитів, які очікують бізнес-менеджери та аналітики. MOLAP-сервер також відповідає на ці запити користувачів,

надсилаючи результати на рівень презентації/користувацький інтерфейс (Dantoni, 2022).

3) Працюючи із цим інструментом, менеджери, аналітики та інші бізнес-професіонали досліджують інформацію, що зберігається в кубах даних на MOLAP-сервері. Вони створюють запити, як правило, за допомогою інструментів «наведи і клацни». Їхні запити надсилаються на MOLAP-сервер для обробки, часто з використанням мови маніпулювання даними, створеної спеціально для роботи з багатовимірними кубами даних і серверами (Dantoni, 2022).

Як працюють між собою всі три рівня даної багатовимірної БД показано нижче (рис. 2.3)

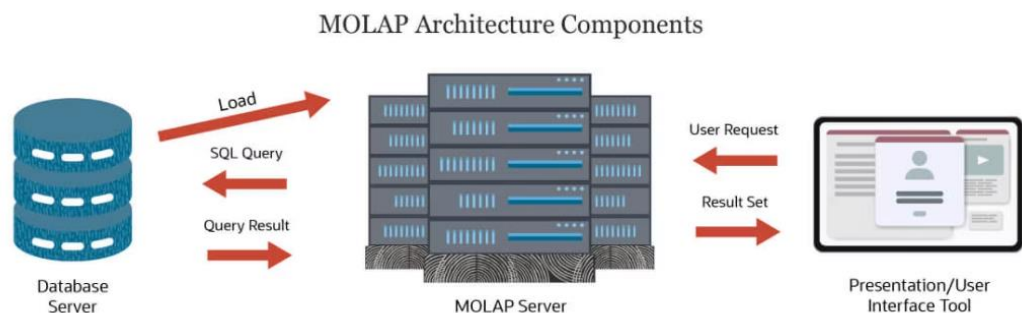


Рис. 2.3 Візуалізація роботи MOLAP системи (Dantoni, 2022).

Для вибору взагалі будь-якої системи потрібне ретельне попереднє проектування та моделювання. Організації, які розглядають можливість впровадження MOLAP, мусять ретельно оцінити свої бізнес-цілі, типи запитань, які вони найчастіше ставлять, доступні їм джерела даних, а також можливості та досвідченість своїх користувачів.

Коли організація розробляє свої куби даних, вона повинна визначити виміри (категорії, такі, як час або місцезнаходження), ієрархії (рівні підкатегорій, які допомагають користувачам згортати або деталізувати свої дані) і показники (оцінювання результатів, таких, як дохід або витрати). Як правило, кожен куб пов'язаний із конкретною сферою аналізу, але куби

можуть бути пов'язані між собою, щоб користувачі могли досліджувати питання, які виходять за межі вузьких організаційних рамок (*Dantoni, 2022*).

Як і будь, яка система MOLAP мають свої переваги та недоліки, їх варто також розглянути. Вважається, що MOLAP має кілька ключових переваг:

- Швидкість виконання запитів. Традиційно головною перевагою MOLAP-систем є швидкість. MOLAP швидша, по-перше, тому, що її архітектура побудована з нуля, щоб відобразити унікальні вимоги аналітичної обробки в обробці кешування, індексування та зберігання, і, по-друге, тому, що ця система виконує більшу частину обчислень заздалегідь. По суті, багато відповідей готові ще до того, як аналітики ставлять запитання (*Dantoni, 2022*).
- Розмір. Куби MOLAP усувають надмірність, тому дані стискаються, вимагаючи менших витрат на зберігання, ніж відповідні дані, що зберігаються в реляційній БД (*Dantoni, 2022*).
- Можливість роботи в автономному режимі. Оскільки запитується куб, а не реляційна БД, яка лежить в основі даних, деяку аналітичну роботу можна виконувати без прямого підключення до реляційної СКБД (*Dantoni, 2022*).

Переваги MOLAP можуть приносити компроміси, які іноді змушують фахівців із баз даних та аналітиків обирати інші рішення OLAP. Традиційно до недоліків MOLAP відносять:

- Обмежена масштабованість і кількість вимірів. Деякі системи MOLAP найкраще працюють з обмеженою кількістю даних і можуть зіткнутися зі значним погіршенням продуктивності при збільшенні обсягів, або кількості вимірів, що аналізуються (*Dantoni, 2022*).
- Потенційно менша гнучкість. MOLAP-системи вимагають ретельного проектування, щоб відобразити типи запитів, які будуть виконуватися. Традиційно, коли потрібні нові типи запитів, виникає потреба в

додатковому професійному проектуванні, що уповільнює прийняття бізнес-рішень (*Dantoni, 2022*).

- Надлишкові дані та затримки. У MOLAP аналітики запитують оброблену копію інформації, що зберігається в окремому наборі кубів, замість того, щоби безпосередньо запитувати оригінальні дані. Відповідно, адміністратори БД мусять керувати двома наборами даних. Це також призводить до затримок: Актуальна інформація, додана до реляційної бази даних, може, бути недоступною для аналітика. Для деяких додатків це не є проблемою, але аналітика все частіше вимагає інформації в режимі реального або близького до реального часу (*Dantoni, 2022*).
- Мовні проблеми. Хоча реляційні бази даних майже всі використовують певний діалект мови SQL для виконання запитів. Куби MOLAP можуть потребувати додаткових мов, для отримання даних і виконання обчислень.

Розглянемо ті інструменти, де використовуються ці підходи. Загалом, інструменти MOLAP — це повний набір архітектурних компонентів (описаних вище), які складають повну реалізацію цієї системи, хоча бізнес-аналітики, які посилаються на даний інструмент, часто мають на увазі рівень представлення, який приймає їхні запити й повертає результати. Було представлено широкий спектр інструментів MOLAP, деякі, з яких можуть підтримувати як MOLAP, так і ROLAP. Зокрема, це: Oracle Essbase, Microsoft SQL Server Analysis Services вони є одними з найвідоміших. Усе частіше MOLAP та інші OLAP-сервіси надаються через хмару, використовуючи властиві хмарі масштабованість і керованість (*Dantoni, 2022*).

2.2 Реляційна аналітична обробка, як вид багатовимірної бази даних

Реляційні OLAP-системи (ROLAP) використовують технологію реляційних баз даних для зберігання даних, а також застосовують спеціалізовані індексні

структури, такі, як бітові індекси, для досягнення високої продуктивності запитів. MOLAP-системи, як правило, забезпечують більш компактне зберігання даних, а також швидший час відгуку на запити (*Pedersen, Jensen, 2001*).

Реляційні OLAP-методи використовують традиційні матеріалізовані подання, багатовимірні OLAP-методи — багатовимірні масиви, графові методи — спеціалізовані графи, а апроксимаційні методи — різні представлення в пам'яті. (*Morfonios, Konakas, Ioannidis, Kotsis, 2007*).

Для того, щоб чітко зрозуміти, чим саме відрізняються моделі OLAP між собою потрібно розуміти, із чого саме складається кожна з них. Модель ROLAP складається з трьох основних компонентів:

1. Сервер БД: Існує на рівні інформації. Складається з даних, які завантажуються на ROLAP-сервер (*Mbaabu, 2021*).

2. Сервер ROLAP: Складається з механізму ROLAP, який існує на рівні додатків (*Mbaabu, 2021*).

3. Інструмент інтерфейсу: Це клієнтський робочий стіл, який існує на рівні презентації (*Mbaabu, 2021*).

Розглянемо, як працює ROLAP. Коли користувач робить запит (складний), сервер ROLAP здобуває дані із сервера СУБД. Потім механізм ROLAP динамічно створює куби даних. Користувач буде переглядати дані з багатовимірної точки (*Mbaabu, 2021*). Приклад роботи даної системи показано нижче (рис. 2.4)

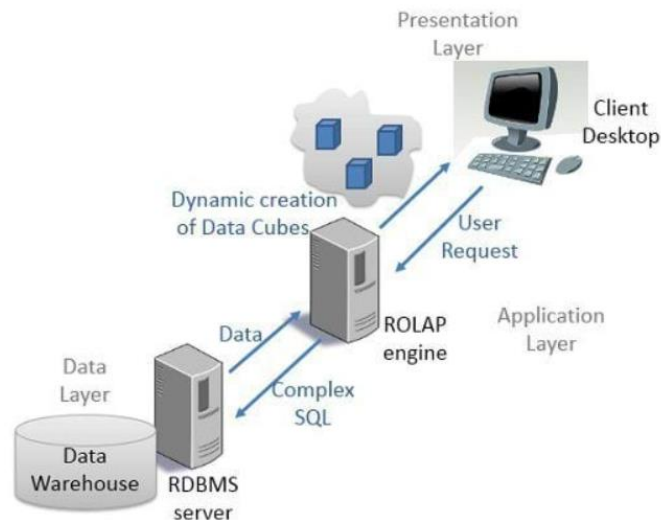


Рис. 2.4 Приклад роботи ROLAP системи (Mbaabu, 2021).

Для кращого розуміння потрібно також розглянути переваги та недоліки таких систем. Для ROLAP виділяють наступні переваги:

- Може, обробляти величезні обсяги даних (Mbaabu, 2021).
- Модель ROLAP, може, ефективно зберігати дані (Mbaabu, 2021).
- Використання реляційної бази даних. Це дозволяє інтегрувати ROLAP-сервер з RDBMS (реляційною системою управління базами даних) (Mbaabu, 2021).

А також виділяють ось такі недоліки:

- Низька продуктивність, особливо при великих обсягах даних (Mbaabu, 2021).
- ROLAP має певні обмеження, пов'язані з SQL. Наприклад, функція SQL має труднощі в обробці складних обчислень (Mbaabu, 2021).

Як було сказано вище існує дві основні схеми реалізації ROLAP. Тому розглянемо їх більш детально окремо одна від одної.

2.2.1 Опис схеми ROLAP типу «Зірка»

Як було сказано вище системи ROLAP використовують звичайні СКБД в якості програмного забезпечення. У зв'язку з цим реалізації таких систем

зазвичай використовують схеми «зірка» (рис. 2.5) або «сніжинка», обидві з яких зберігають дані в таблицях фактів і таблицях розмірностей. Оскільки саме такі моделі допомагають реалізувати багатовимірність. В даних схемах таблиця фактів містить один рядок для кожного факту в кубі. Вона має стовпчик для кожного показника, що містить значення показника для конкретного факту, а також стовпчик для кожного виміру, який містить зовнішній ключ, що посилається на таблицю вимірів для конкретного виміру (Pedersen, Jensen, 2001).

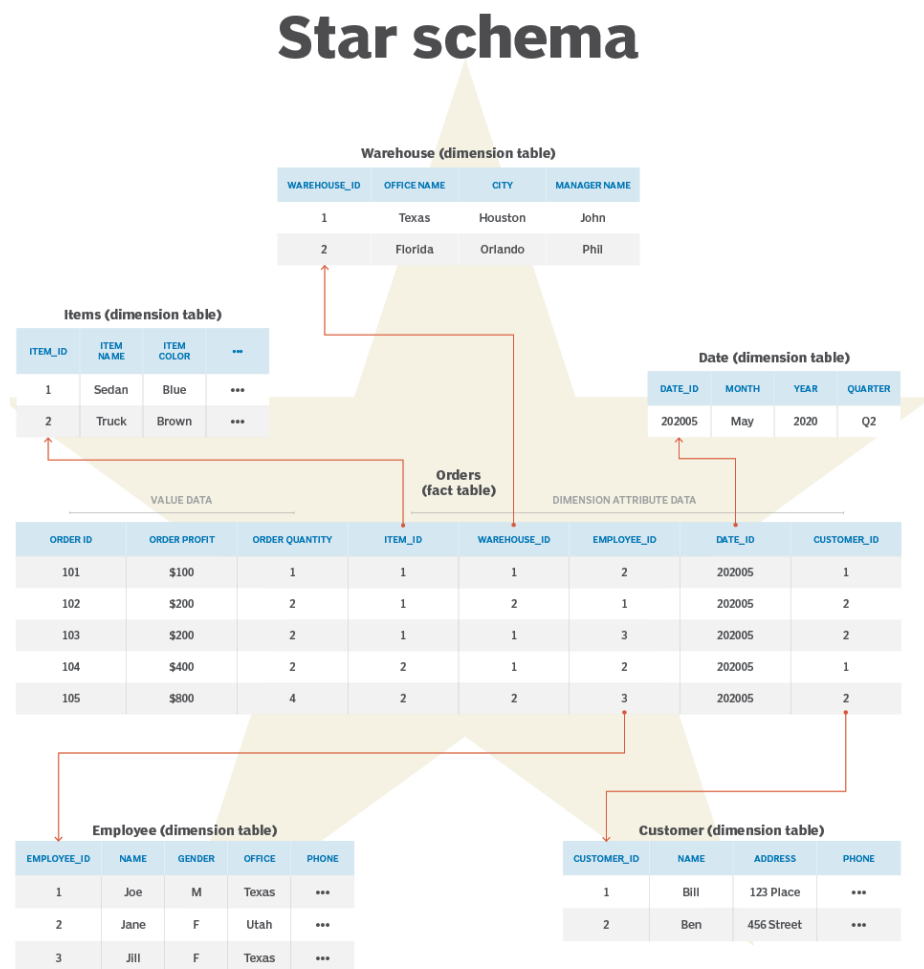


Рис.2.5 Приклад схеми по типу «зірка» TechTarget Contributor (2021).

Ці дві схеми відрізняються тим, як вони обробляють виміри, і вибір між ними значною мірою залежить від бажаних властивостей системи, що розробляється. Як показано на рисунку 2.5, схема "зірка" має одну таблицю для кожного розміру. Таблиця вимірів містить ключовий стовпчик, один

стовпчик для кожного рівня виміру, що містить текстовий опис значень цього рівня, і один стовпчик для кожної властивості рівня у вимірі. Таблиця фактів зіркоподібної схеми містить ціну продажу для одного конкретного продажу та пов'язані з нею значення розмірності. Вона має стовпчик зовнішнього ключа для кожного з трьох вимірів: товар, місце і час.

Таблиці вимірів мають відповідні ключові стовпчики і по одному стовпчику для кожного рівня виміру - наприклад, ID Товару, Ім'я товару, колір товару. Ключовий стовпчик таблиці розмірностей зазвичай є фіктивним цілочисельним ключем без будь-якої семантики. Це запобігає неправильному використанню ключів, забезпечує краще використання сховища і надає більшу підтримку для оновлення розмірностей, ніж інформативні ключі з вихідних систем (*Pedersen, Jensen, 2001*).

Надлишковість буде мати місце в даних вищих рівнів. Наприклад, оскільки травень будь-якого року має 31 день, відповідно через це значення року "2020" буде повторено 31 раз в межах даної таблиці. Оскільки розмірності зазвичай займають, лише від одного до п'яти відсотків загального обсягу пам'яті куба, надлишковість не є проблемою зберігання. Крім того, централізована обробка оновлень розмірностей забезпечує узгодженість. Таким чином, використання денормалізованих таблиць розмірностей, які підтримують простіші формулювання більш ефективних запитів, часто є вигідним (*Pedersen, Jensen, 2001*).

2.2.2. Опис схеми ROLAP типу «Сніжинка»

Дана схема ROLAP містить одну таблицю для кожного рівня розмірності, щоб уникнути надмірності, що, може, бути корисним у деяких ситуаціях. Кожна таблиця розмірності містить ключ, стовпчик з текстовим описом значень рівня і, можливо, стовпчики для властивостей рівня. Таблиці нижчих

рівнів також містять зовнішній ключ до рівня, що їх містить. Наприклад, таблиця предметів містить ID з іншої таблиці кольорів, яка більш детально описує колір це відтінок металік, чи глянець. Відповідно одне значення кольору червоний, може, повторюватися з іншим атрибутом відтинку (Pedersen, Jensen, 2001). Схему типу "сніжинка" показано на рис. (2.6).

Snowflake schema

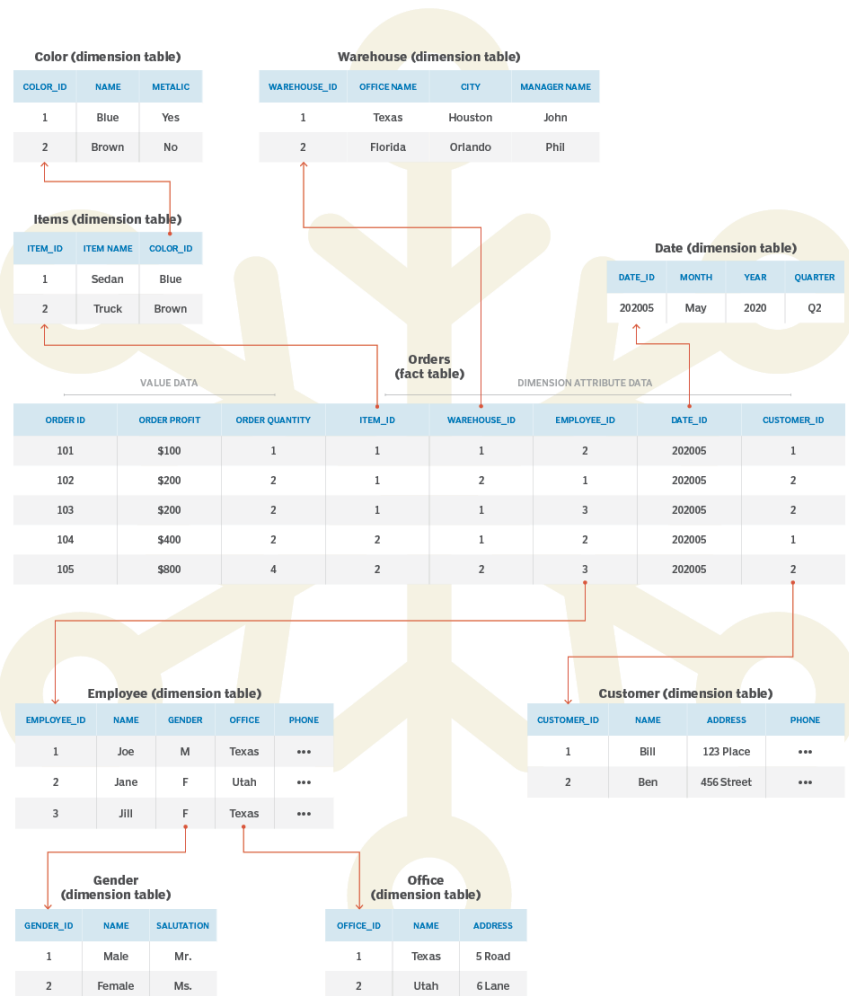


Рис.2.6 Приклад схеми по типу «сніжинка» TechTarget Contributor (2021).

Оскільки схеми все ж таки відрізняються варто розглянути ключові особливості, а також переваги та недоліки такої схеми. Розмірна модель сніжинки має наступні особливості:

- Схеми сніжинки займає мало місця на диску (Snowflake Schema in Data Warehouse Model - GeeksforGeeks, 2023).

- Легко реалізувати розмірність, яка додається до схеми (*Snowflake Schema in Data Warehouse Model - GeeksforGeeks, 2023*).
- Існує декілька таблиць, тому продуктивність знижується (*Snowflake Schema in Data Warehouse Model - GeeksforGeeks, 2023*).
- Таблиця розмірностей складається з двох або більше наборів атрибутів, які визначають інформацію на різних зернах (*Snowflake Schema in Data Warehouse Model - GeeksforGeeks, 2023*).
- Набори атрибутів однієї і тієї ж розмірності заповнюються різними системами-джерелами (*Snowflake Schema in Data Warehouse Model - GeeksforGeeks, 2023*).

До переваг відносять:

- Забезпечення структурованості дані, що зменшує проблему їх цілісності (*Snowflake Schema in Data Warehouse Model - GeeksforGeeks, 2023*).
- Використання мало місця на диску, оскільки дані добре структуровані (*Snowflake Schema in Data Warehouse Model - GeeksforGeeks, 2023*).

До недоліків відносять:

- Сніжинка зменшує простір, який займають розмірні таблиці, але порівняно з усім сховищем даних економія зазвичай незначна (*Snowflake Schema in Data Warehouse Model - GeeksforGeeks, 2023*).
- Кілька ієрархій, які можуть належати до одного розміру, повинні бути спроектовані з мінімально можливою деталізацією (*Snowflake Schema in Data Warehouse Model - GeeksforGeeks, 2023*).

Як згадувалось вище існує третій різновид онлайн-аналітики це поєднання двох попередніх видів в одну систему. А саме поєднання MOLAP та ROLAP таке з'єднання називають HOLAP – що так і перекладається гібридна онлайн-аналітика.

2.3 Гібридна аналітична обробка, як вид багатовимірної бази даних.

HOLAP розшифровується як Hybrid Online Analytical Processing, що в перекладі означає гібридна онлайн-аналітична обробка. Як випливає з назви, це гібридна модель, що складається з двох різних типів OLAP, таких, як реляційна онлайн-аналітична обробка (ROLAP) і багатовимірна онлайн-аналітична обробка (MOLAP). OLAP — це не що інше, як метод, що застосовується для бізнес-аналітики, наприклад, для управління та аналізу даних у системах їх сховищ. У моделі, що ґрунтується на HOLAP, система складається з таблиць і БД, які побудовані за допомогою реляційних і багатовимірних систем управління базами даних. Іншими словами, HOLAP дає змогу зберігати дані, як у реляційних, так і в багатовимірних базах даних, використовуючи функцію кешування для багаторазово використовуваного набору даних і складних запитів для багатовимірних таблиць (Pedamkar, б. д.).

Розглянемо детально структуру такого типу системи. У системі HOLAP користувач, може, отримати доступ до даних із реляційної та багатовимірної БД, використовуючи прості запити для і складні агреговані запити для багатовимірної бази даних. Оскільки процес створення запитів є складним і ті самі запити можуть повторно використовуватися в багатовимірній системі, ефективно побудувати кешований сервер тільки для агрегованих і багаторазово отриманих даних із реляційної БД (Pedamkar, б. д.). Для кращого розуміння розглянемо візуалізацію такої системи, яку показано на (рис. 2.7).

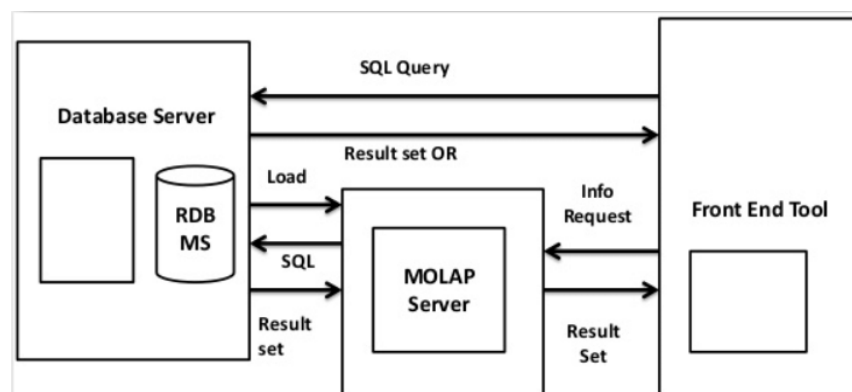


Рис. 2.7 Візуалізація роботи системи HOLAP (Мбаабу, 2021).

Користувач, тобто аналітик або фахівець зі звітності, робить запит до HОLAP для отримання необхідної інформації. Запити стосуються, як реляційної БД, так і багатовимірної бази даних. Прості запити до реляційної БД передають дані до системи кешування, а складні запити зберігаються в системі кешування разом із даними, запитаними для поточного потоку. Дані, відновлені з обох систем, потім надаються користувачеві для аналізу або створення звітів для прийняття бізнес-рішень (*Pedamkar, б. д.*).

Оскільки багатовимірна система є складною, кешування потрібно підтримувати постійно, тобто кожного разу, коли виконується оновлення в системі багатовимірної бази даних. СКБД не потребує кешування через свою просту структуру. Вона використовує систему кешування для об'єднання результатів із результатами, отриманими з багатовимірної бази даних. Ця модель дає змогу потоку даних бути більш ефективним, оскільки реляційна БД зосереджена на кількості даних, а багатовимірна — на оцінюванні даних. Коли агрегати кешуються, запити зберігаються, щоб не витратити зусилля на створення і застосування одних і тих же запитів повторно (*Pedamkar, б. д.*).

Це допомагає підвищити продуктивність системи HОLAP, займаючи при цьому величезний обсяг пам'яті. Оскільки HОLAP, може, мати переваги як ROLAP, так і MOLAP, жертвувати простором для зберігання даних доводиться (*Pedamkar, б. д.*).

Для даного типу СППР виділяють наступні недоліки та переваги. Перевагами є:

- Краща продуктивність і масштабованість, оскільки поєднує в собі багатовимірні та реляційні атрибути аналітичної обробки в режимі онлайн (*Mbaabu, 2021*).

- Винахідливий інструмент аналітичної обробки, якщо ми очікуємо, що розмір даних збільшиться (*Mbaabu, 2021*).

· Його здатність до обробки вища, ніж у двох інших інструментів аналітичної обробки (Mbaabu, 2021).

Недоліками є:

· Використання величезного обсягу пам'яті, оскільки складається з даних із двох БД (Mbaabu, 2021).

· Модель потребує частих оновлень через свою складну природу (Mbaabu, 2021).

Якщо звести всі особливості кожної з трьох типів систем, то вона буде мати такий вигляд: (табл. 2.2)

Таблиця 2.2 Порівняння типів СППР (Mbaabu, 2021).

Назва	MOLAP	ROLAP	HOLAP
Значення	Багатовимірна онлайн-аналітична обробка	Реляційна онлайн-аналітична обробка	Гібридна онлайн-аналітична обробка
Зберігання даних	Він зберігає дані в багатовимірній базі даних.	Він зберігає дані в реляційній базі даних.	Він зберігає дані в реляційній базі даних та в багатовимірній
Технічна реалізація	Використовує техніку розрідженої матриці.	Використовує мову структурованих запитів (SQL).	Використовує комбінацію SQL і техніки розріджених матриць.
Обсяг даних	Може обробляти обмежений обсяг даних.	Обробляє величезні дані.	Може обробляти величезні обсяги даних.
Розрахунковий вигляд	Багатовимірне зображення є статичним.	Багатовимірне зображення є динамічним.	Багатовимірне зображення є динамічним.
Упорядкування даних	Впорядковує дані в кубах даних.	Впорядковує дані в рядки і стовпці (таблиці).	Існує багатовимірне розташування даних

У цьому розділі детально розглянуто моделі реалізації багатовимірних баз даних, а саме MOLAP, HOLAP, а також схеми, які найчастіше використовуються в ROLAP. Описано всі їх недоліки та переваги.

На основі пропонованого аналізу були зроблені висновки про те, який саме вид багатовимірної бази даних краще обрати для даних електророзвідки. Було встановлено, що варто обрати реалізації ROLAP, а саме схему сніжинки. Оскільки такий тип дає великі можливості, та є простим у реалізації та підтримці. Також використовуючи даний тип СППР можна застосувати СКБД, у якій дуже зручно зберігати просторові дані, а також їх аналізувати та візуалізувати.

3. ПРАКТИЧНА ЧАСТИНА

В якості вхідних даних було взято файли виконання вимірювання методом електрорезистивної томографії.

Файли мають розширення .dat. Містять не дуже багато інформації. На мою думку, даної інформації взагалі не достатньо для створення повноцінної багатовимірної бази даних. Тому решта інформації, які я вважаю необхідно також врахувати при аналізі електричного опору отриманого в результаті вимірювання були згенеровані випадковим чином. Для чіткого розуміння потрібно розглянути структуру файлу. Таблиця 3.1. Демонструє фрагмент даного файлу:

Таблиця 3.1 фрагмент вхідного файлу.

Wenner-Schlumberger survey			
1.00			
7			
1219			
1			
0			
X -координата	«Н» фактор	значення «а» фактору відстані	Виміряний опір
1.5	1	1	19.9528
2.5	1	1	27.8069
3.5	1	1	90.1845
4.5	1	1	129.0909
5.5	1	1	177.5917
6.5	1	1	263.7877
7.5	1	1	230.6737
8.5	1	1	312.7889
9.5	1	1	292.4173
10.5	1	1	201.7366

Перші шість рядків є так званою шапкою. Кожен рядок несе чітко свою інформацію:

1. Назва площі, або назва файлу.
2. Відстані між електродами.
3. Array-number це тип вимірювання.

4. Загальна кількість точок в даному файлі.
5. Для координат; Тип х-локації для точок даних. Стоїть 0, якщо для вказівки х-позиції використовується розташування першого електрода в масиві. Стоїть 1, якщо використовується середина масиву (тобто розташування точки даних у псевдорозрізі).
6. Стоїть 0 для даних без IP; Використовується 1 для даних з IP. IP – включаючи поляризацію.
7. В кінці файлу вставляється 4 нулі в кожному рядку 0 це означає кінець файлу.

Після так званої шапки йде розбиття на 4 стовпці їх розділяють пробіли. Їх варто теж описати та розглянути більш детально:

1. X -координата;
2. Значення «а» відстані вимірюється при методі диполь-диполь.
3. «Н» фактор. Використовується для методу Диполь-диполь. Вимірюється в значенні від 0.5 до 3.5
4. Виміряне значення стійкості

Координата X в даному файлі є умовною. Але з її допомогою можна прослідкувати зміни профіля зйомки, а саме зближення електродів до центру.

Фактори «а» та «н» використовують у методі диполь-диполь. Для диполь-дипольної конфігурації масивів простір між парою електродів А, М та електродами В, М розглядається, як «а» простір між парою електродів М, N розглядається як n.a (рис. 3.1). Для збільшення глибини дослідження, зберігаючи відстань постійною і збільшуємо відстань n.a. у коефіцієнті між "1 і 6". Для другого вимірювання крок а збільшується на "2а" тих самих значень "n". Цей процес повторюється для наступних вимірювань (Loke, 2004).

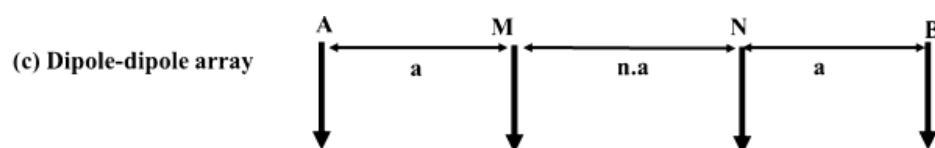


Рис. 3.1 Схематичне зображення конфігурації електродів для методу диполь-диполь (Singh & Sharma, 2022)

Для кращого розуміння, що саме потрібно моделювати та, які варто виділяти розмірності в багатовимірні базі даних варто дуже детально ознайомитися з методикою проведення вимірювання методом електрорезистивної томографії.

3.1 Огляд методу електрорезистивна томографії.

Електрорезистивна томографія — популярний метод дослідження неглибоких шарів ґрунту, який використовується для вирішення гідрогеологічних, інженерних та сільськогосподарських завдань. Він, може, застосовуватися в широкому діапазоні масштабів (*Gunther & Rucker, 2012*).

У даному методі вимірювання використовують питомий опір постійного струму (DC) — традиційний геофізичний метод. Він використовує два електроди для подачі електричного струму в землю, а інші два електроди — для вимірювання різниці електричних потенціалів. Вимірювання часто проводять уздовж лінії або в певній ділянці земної поверхні, а потім спостережувані різниці потенціалів перетворюють на криві зондування або псевдоперерізи уявного питомого опору, які вказують на зміни питомого опору підземних гірських порід. Аналіз цих даних дає змогу знайти підземні аномалії питомого опору або окреслити підземну структуру природного чи штучного походження. Точне походження важко встановити, для чого проводять інші дослідження, як поверхневі так і буріння свердловин у місце аномалії (*Zhou, 2019*).

З розвитком комп'ютерних технологій та чисельних обчислювальних методів стало можливим точне чисельне моделювання підземного електричного поля та отримання великої кількості даних, тому традиційна розвідка питомого електричного опору постійним струмом була розвинута до комп'ютеризованого методу геотомографії, який називається електрична

томографія питомого електричного опору (ERT), що використовує багатоелектродне обладнання або систему для автоматичного отримання великої кількості інформації та застосовує комп'ютерне програмне забезпечення для реконструкції структури питомого електричного опору підземних порід на основі спостережених даних (Zhou, 2019).

Варто також розглянути фундаментальну теорію для цього методу вимірювання опору. Згідно з неперервністю електричного струму, виконується наступне інтегральне рівняння виконується в будь-якій точці провідного середовища (Zhou, 2019). Це відображає формула (3.1):

$$\oint_{\Gamma} \mathbf{J} \cdot d\Gamma = -I \quad (3.1.)$$

де Γ - повна або напівсферична поверхня, яка охоплює електрод, що інжектує електричний струм J з силою I (рис. 3.2). Відповідно до закону Ома впливає наступне (Zhou, 2019). формула (3.2):

$$\mathbf{J} = \sigma \mathbf{E} = -\sigma \nabla U, \quad (3.2)$$

Та властивість дельта функції $\delta(x-x_s)$, з формули 3.1. дає наступне (Zhou, 2019): формула (3.3)

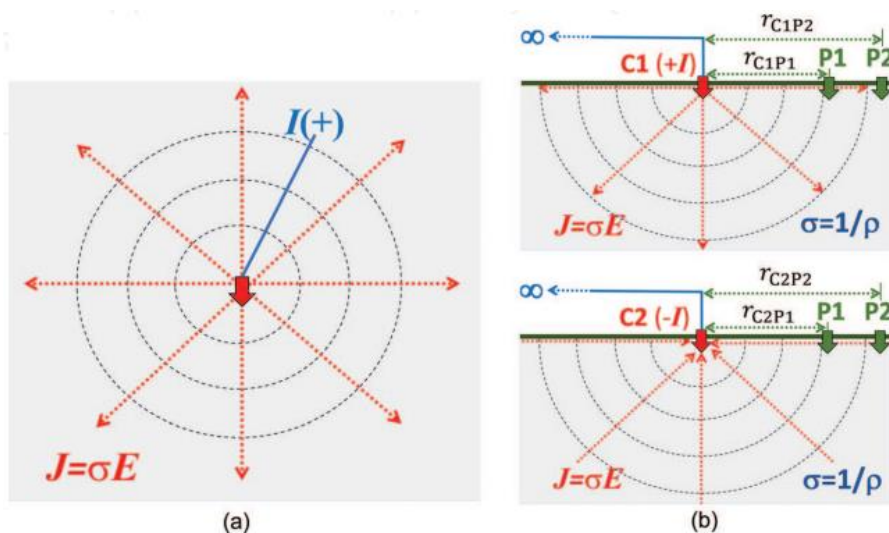


Рис. 3.2. Густина електричного струму J навколо електрода в (а) повному просторі та (б) напівпросторі (Zhou, 2019).

$$-\oint_{\Gamma} \sigma \nabla U \cdot d\Gamma = -I \iiint_{\Omega} \delta(x-x_s) d\Omega \quad (3.3)$$

де σ - провідність середовища, U - електричний потенціал, Ω - об'єм сферичної поверхні Γ , and $x_s = (x_s, y_s, z_s)$ це місце розташування струмового електрода. Застосування теореми про дивергенцію до рівняння (3.3) приводить до наступного (Zhou, 2019): формула (3.4)

$$\iiint_{\Omega} [\nabla \cdot (\sigma \nabla U) + I \delta(\mathbf{x} - \mathbf{x}_s)] = 0 \quad (3.4)$$

Варто зазначити, що рівняння (3.4) виконується скрізь у середовищі, так що рівняння електричного поля формула (3.5):

$$\nabla \cdot (\sigma \nabla U) = -I \delta(\mathbf{x} - \mathbf{x}_s), \mathbf{x} \in \Omega \quad (3.5)$$

У загальному випадку провідність $\sigma(\mathbf{x}) = \sigma(x, y, z)$ змінюється за трьома координатами так, що вона визначає тривимірну геологічну модель, а електричний потенціал $U = U(x, y, z)$ стає тривимірною функцією заданого нагнітання струму $I \delta(\mathbf{x} - \mathbf{x}_s)$ і провідності $\sigma(\mathbf{x})$ так, як це показано на (рис. 3.3 а). Якщо провідність є константою координати y , тобто $\sigma(\mathbf{x}) = \sigma(x, z)$, то вона визначає 2D геологічну модель (рис. 3.3. б), але електричний потенціал $U(x, y, z)$ все ще є 3D функцією координат через точкове нагнітання струму $I \delta(\mathbf{x} - \mathbf{x}_s)$. Застосовуючи косинусне перетворення Фур'є $F_c\{*\}$ до координати y рівняння (3.5), (Zhou, 2019) керуюче рівняння змінюється на формулу (3.6):

$$\nabla \cdot (\sigma \nabla \tilde{U}) + k_y^2 \tilde{U} = -I \delta(\mathbf{x} - \mathbf{x}_s) / 2, \mathbf{x} \in \Omega, \quad (3.6)$$

яке часто називають 2.5D керуючим рівнянням. Тут k_y - хвильове число, і $\tilde{U}(x, k_y, z) = F_c\{U(x, y, z)\}$ стає спектром електричного потенціалу в області хвильових чисел. Координатний вектор $\mathbf{x} = (x, z)$ і градієнт $\nabla = (\partial_x, \partial_z)$ є двовимірними версіями. Якщо задано геологічну модель $\sigma(x, y, z)$ або $\sigma(x, z)$ і нагнітання струму $I \delta(\mathbf{x} - \mathbf{x}_s)$, то, розв'язавши визначальне рівняння (3.5) або (3.6), отримуємо електричний потенціал $U(x, y, z)$ для тривимірної моделі провідності або його спектр $\tilde{U}(x, k_y, z)$ для двовимірної моделі провідності, а потім виконують обернене косинусне перетворення Фур'є $U(x, y, z) = F_c^{-1}\{\tilde{U}(x, k_y, z)\}$ для отримання електричного потенціалу $U(x, y, z)$. Ці обчислення називаються

прямим моделюванням. У теоретичних розрахунках або чисельному прямому моделюванні часто застосовують функцію Гріна G або \tilde{G} - відгук електричного потенціалу на введення одиничного струму ($I = 1$ A) - для того, щоб вилучити величину електричного струму I з рівнянь (3.5) і (3.6). Можна зосередитися на обчисленні функції Гріна, і тоді електричний потенціал U або \tilde{U} обчислюється множенням функції Гріна G на величину практичного електричного струму I , тобто $U = I G$ (Zhou, 2019).

Найпростішою геологічною моделлю є однорідний півпростір. Застосовуючи рівняння (3.3) (еквівалентне рівнянню (3.5)) до постійного середовища, поверхневий інтеграл обчислюється за формулою (Zhou, 2019). (3.7):

$$\oint_{\Gamma} \sigma \nabla U \cdot d\Gamma = \sigma \nabla U (2\pi r^2) = -I \quad (3.7)$$

Яка дає електричний потенціал на відстані r від джерела струму, яке описане у формулі (3.8):

$$U = \frac{I}{2\pi\sigma r} = \frac{\rho I}{2\pi r} \quad (3.8)$$

Де $\rho = 1/\sigma$ - питомий опір. На практиці, щоб ввести електричний струм в землю, необхідно використовувати пару струмових електродів $C1$ і $C2$, один з яких є позитивним (+), а інший - негативним (-). інший - негативний (-) (Zhou, 2019).

Відмінність між 2D та 3D моделлю можна побачити на (рис. 3.3)

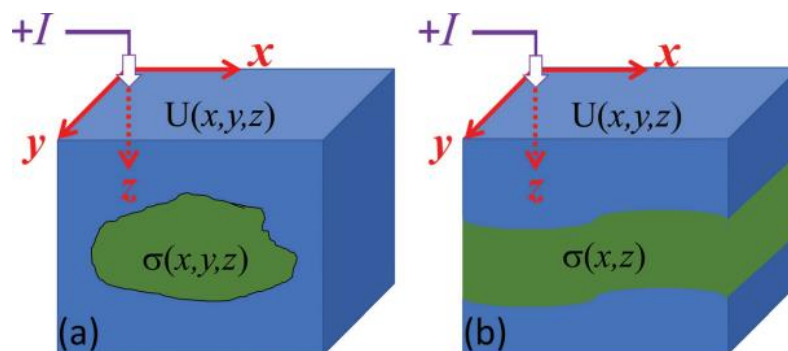


Рис.3.3 Зображення (а) 2D та (б) 3D геологічної моделі, визначеної провідністю $\sigma(x,z)$ та $\sigma(x,y,z)$ відповідно (Zhou, 2019).

Для того, щоб отримати видимий питомий опір у полях, було розроблено багато електродних решіток у традиційній розвідці питомого опору постійного струму. Від ERT всі сфери вимагають високої щільності даних і хорошого покриття земної поверхні для отримання зображень підземних об'єктів із високою роздільною здатністю. Далін і Чжоу (2004) провели синтетичні експерименти ERT із використанням 10 різних електричних решіток і порівняли результати їх візуалізації для чотирьох геологічних моделей: заглибленого каналу, вузької дамби, занурених блоків і ставків для відходів. Вони продемонстрували, що дві триелектродні решітки (полюс-диполь і диполь-полюс) і три чотириелектродні решітки (диполь-диполь, решітка Schlumberger і градієнтна решітка) дають задовільні зображення підземних об'єктів. Однак через використання віддалених електродів трьохелектродні решітки рідко застосовуються для ERT на практиці, тому популярними стають чотириелектродні решітки. Зокрема, градієнтна решітка добре підходить для багатоканального збору даних і, може, значно підвищити швидкість збору даних у польових умовах, і водночас дає вищу щільність даних і меншу чутливість до шумів, ніж дипольно-дипольна решітка. На рисунку 3.4 показано три чотириелектродні решітки (верхній ряд) та їхні псевдоперерізи точок даних x_{pa} , z_{pa} (середній ряд) і їхні геометричні коефіцієнти (нижній ряд) для схеми розміщення 81 електрода на земній поверхні. На рисунку 3.4 показано, що диполь-дипольна решітка має найбільший діапазон геометричного фактора серед цих трьохелектродних решіток, а градієнтна решітка має найменший діапазон, тому вона добре підходить для збору інформацію із високою щільністю даних і гарною якістю (Dahlin, Zhou, 2004).

Дивлячись на Рисунок 3.4, можна побачити, що градієнтна решітка використовує для вимірювання різниці потенціалів, лише потенційні електроди між C1 і C2. Він не виконує вимірювання на потенційних електродах

за межами C1 і C2. Це означає, що градієнтна матриця фактично пропускає зовнішні дані для кожної пари струмових електродів (Zhou, 2019).

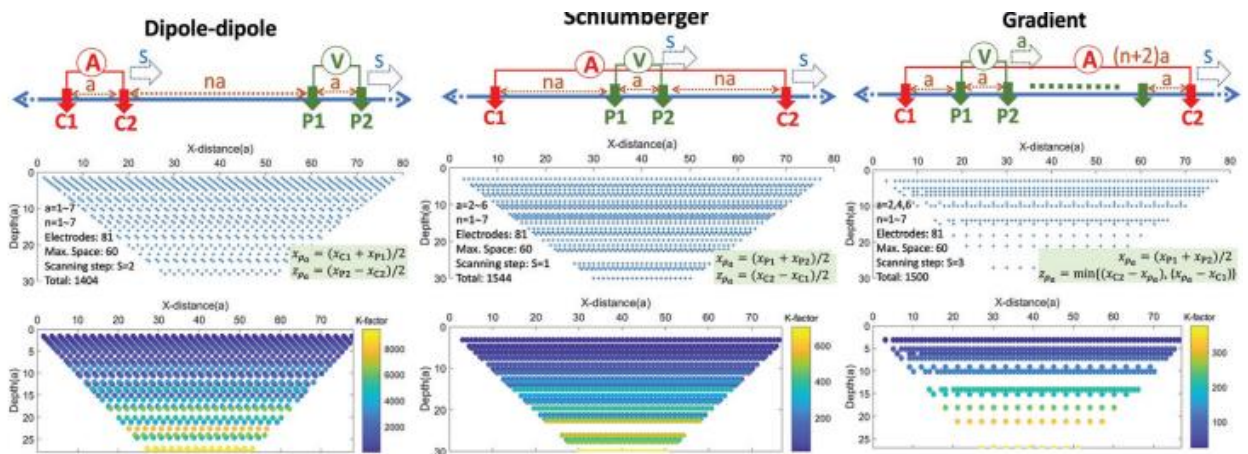


Рис. 3.4 Диполь-дипольні, Шлюмберже та градієнтні масиви (верхній ряд) для збору даних ERT та приклади їх псевдорозрізи точок даних (x_{pa} , z_{pa}) (середній ряд) та геометричні коефіцієнти (нижній ряд) (Zhou, 2019).

В рамках даної роботи були в якості вхідних даних були взяті ERT виконані методом Диполь-Диполь. Тепер варто розглянути, які були виділені розмірності та обґрунтування їх необхідності.

3.2 Розробка багатовимірної бази даних

У рамках цієї роботи розроблено головну частину системи прийняття рішень типу ROLAP, а саме в рамках реляційної бази даних, була реалізовано багатовимірна схема типу «сніжинка». Вона є головною оскільки з її допомогою можна збирати, зберігати й потім обробляти вхідні масиви даних.

У рамках 1 фази розроблення СППР є розвідки, збір інформації та пошук проблем. Проблема була сформована таким чином. Повна відсутність прикладів використання систем типу OLAP для даних ERT, та багатовимірних БД, зокрема. Як вхідні дані були використані результати польових вимірювань декількох площ.

На другому етапі проектування альтернатив були вирішено розробити таку бази даних, яка суттєво розкриває абсолютно всі аспекти, які можуть давати вплив та бути причиною похибок вимірювання електрорезистивної томографії. У рамках другої фази було розглянуто метод реалізації СППР як об'єднання ГІС та БД, де шаруватість, тобто виокремлення показників один від одного за допомогою карт. У межах даної роботи це, можливо, оскільки більшу половину методів не можливо візуалізувати на мапі. Також було розглянуто побудову багатовимірної бази даних, яка дає можливість досягнути поставлену ціль.

На третьому етапі проектування рішення було обрано систему керування базами даних PostgreSQL. Оскільки вона має дуже гарну підтримку просторових даних. Зокрема, має наступні типи (табл 3.2):

Таблиця 3.2 Просторові типи даних (Geometric Types. (б. д.).

Ім'я	Розмір сховища	Розшифрування	Приклад даних
Точка	16 байт	Точка на площині	(x,y)
Лінія	32 байти	Нескінченна лінія	{A,B,C}
Лінійний сегмент	32 байти	Скінченний відрізок прямої	((x1,y1),(x2,y2))
Коробка	32 байти	Прямокутна коробка	((x1,y1),(x2,y2))
Шлях	16+16n байтів	Замкнутий шлях (схожий на багатокутник)	((x1,y1),...)
Шлях	16+16n байтів	Відкритий шлях	[(x1,y1),...]
Полігон	40+16n байтів	Багатокутник (подібно до замкненого контуру)	((x1,y1),...)
Коло	24 байти	Коло	<(x,y),r> (центральна точка та радіус)

Як видно, з таблиця ця СКБД налічує 8 різних типів просторових даних. Що є перевагою для наших цілей. Також варто відмітити наступні переваги для даної СКБД:

1. PostgreSQL надає широкий спектр розширених типів даних, включаючи масиви, геометричні типи, про які вже згадано та повнотекстовий пошук, які є корисними для даних ERT.

2. Масштабованість: PostgreSQL має високу масштабованість і, може, легко обробляти великі обсяги даних. Вона розроблена для роботи на системах із декількома процесорами й, може, використовуватися в кластерному середовищі для забезпечення ще більшої масштабованості.

3. Продуктивність: PostgreSQL відома своєю продуктивністю і, може, швидко обробляти складні запити до великих наборів даних. Вона має ефективний оптимізатор запитів і підтримує різні методи індексування, які можна використовувати для прискорення запитів.

4. Відкритий вихідний код: PostgreSQL — це система баз даних із відкритим вихідним кодом, яка є вільною у використанні та поширенні. Це робить її ідеальним вибором для зберігання будь-яких даних, оскільки вона, може, допомогти зменшити витрати й забезпечити гнучкість для організацій з обмеженим бюджетом. Крім того, відкритий вихідний код означає, що існує велика спільнота розробників, які роблять свій внесок у її розвиток і надають підтримку користувачам.

5. PostGIS — це розширення PostgreSQL, яке надає можливості просторової бази даних, включаючи підтримку типів і функцій геопросторових даних. Це робить його потужним інструментом для візуалізації даних у геопросторовому контексті. За допомогою PostGIS можна створювати карти, виконувати просторові запити і візуалізувати дані.

Саме створення багатовимірної бази даних є оптимальним не тільки для просторових даних, хоча і для них також, а і для зручного виокремлення кожного аспекту, який, може, впливати на кінцевий результат. А також аналіз усіх цих аспектів, як окремо один від одного так і в різних комбінація об'єднання. Таким чином була розроблена наступна схема (табл. 3.3)

Таблиця 3.3 Схема бази даних

	Розмірності	Підрозмірності
<p>Таблиця фактів:</p> <ul style="list-style-type: none"> • ID Вимірювання • «А» фактор • «Н» фактор • Виміряний опір • Загальна кількість точок • IP поляризації • Тип координат • Розташування • Висота • Глибина проникнення • ID площі • ID приладу • ID Хронології • ID Оператора • ID Погоди • ID Літології • ID Підземних вод 	<p>Таблиця площ:</p> <ul style="list-style-type: none"> • Координати • Система координат • Назва • Номер дозволу • Площа • Середня висота 	
	<p>Тип вимірювання (Array Number):</p> <ul style="list-style-type: none"> • Номер • Повна назва 	
	<p>Хронологічна шкала:</p> <ul style="list-style-type: none"> • Еон • Ера • Період • Епоха • Вік 	
	<p>Підземні води:</p> <ul style="list-style-type: none"> • Номер водосного горизонту • Потужність • Глибина початку 	
	<p>Літологія:</p> <ul style="list-style-type: none"> • Назва породи • Середня потужність • Номер підтипу • Середній опір 	<p>Літологічні підтипи:</p> <ul style="list-style-type: none"> • Назва • Максимальний опір • Мінімальний опір • Середнє значення • Мінеральний склад • Пористість
	<p>Вимірювальний пристрій:</p> <ul style="list-style-type: none"> • Назва • Кількість діапазонів • Частота • Вольтаж • Точність • Тип вимірювання • Матеріали дротів 	
	<p>Оператор:</p> <ul style="list-style-type: none"> • Прізвище та Ім'я • Організація • Позиція • Досвід (в роках) 	
	<p>Погода:</p> <ul style="list-style-type: none"> • Кількість опадів за день, тиждень, місяць до зйомки • Середня температура за день, тиждень, місяць до зйомки • Загальний тип клімату 	

	Агрогрупи: <ul style="list-style-type: none"> • Шифр • Повна назва • Вміст гумусу % • Кислотність % • Засоленість % • Пористість % • Питома маса • Об'ємна маса • Потужність м. • Координата шурфа 	
--	--	--

Для кращого розуміння варто розібрати причину виділення та типи даних у кожній розмірності представленої БД. Зв'язки між таблицями розмірностей та таблицею фактів, або між розмірностями ієрархії, як слідує із самої логіки багатовимірної бази даних багато до одного, де багато даних, може, міститися в таблиці фактів, а в таблиці розмірностей вони є унікальними й повинні відповідати першій нормальній формі. Такі зв'язки забезпечені тим, що в таблицях розмірностей існує атрибут ID, який просто дає порядковий номер для кожного рядка даних він у свою чергу виступає первинним ключем (primary key). Який використовується в якості зовнішнього ключа в таблиці фактів, або в наступній розмірності.

Усього розроблена в рамках цієї роботи схема має 11 таблиць, 9 з них це таблиці розмірностей 1 з яких є розгалуженням іншої. Відгалуження в схемі сніжинки припустимі лише тоді, коли дотримується ієрархізація даних. В даній схемі це можна й необхідно зробити для розмірності літологія, оскільки вони у свою чергу залежить від підтипів гірських порід. Усі ці розмірності це те від чого, залежить виміряні значення опору. І відповідно одинадцята це таблиця фактів.

Важко сперечатися з тим, що питомий електричний опір різних типів гірських порід, може, сильно відрізнятись через відмінності в пористості, мінеральному складі та вмісті води. Наприклад, значення питомого електричного опору для пісковиків і вапняків зазвичай вищі, ніж для сланців і глин, які мають вищу пористість і менший вміст мінералів. Також існує

зворотня залежність низькі показники опору характерні для глин, а високі для пісковиків. Розуміння цих літологічних залежностей має вирішальне значення для точної інтерпретації даних про питомий електричний опір і виявлення підповерхневих геологічних особливостей. Приклад на скільки сильно опір відрізняється залежно від порід показано на рис. 3.5

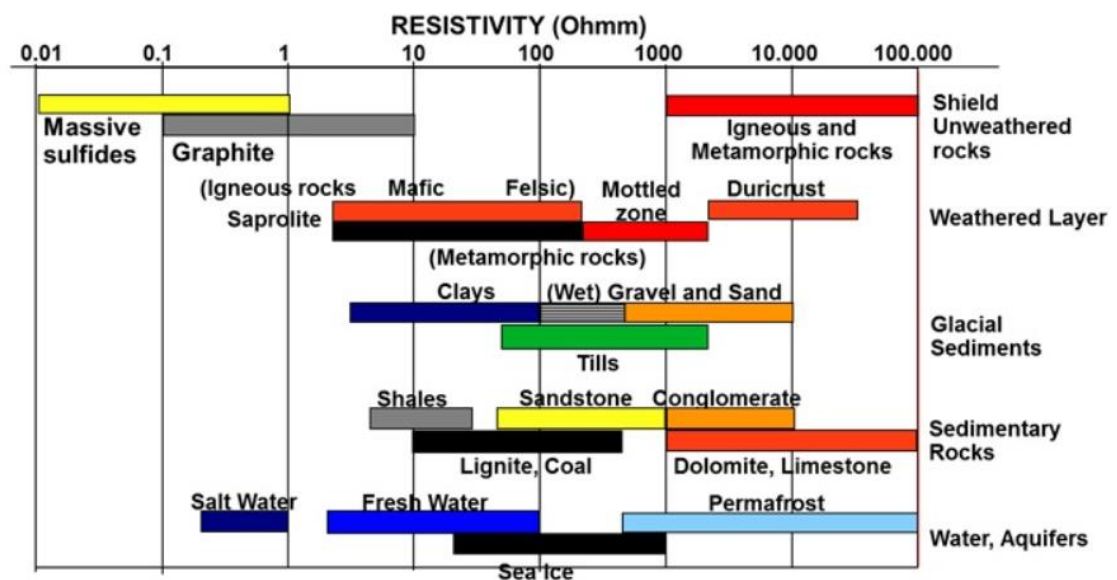


Рис. 3.5 Залежність опору від типу порід (What is VES, ERT, IP and SP Electrical Resistivity ? - Seis Tech, б. д.)

Були виділено розмірність з підтипами порід оскільки порода з однією і тією самою назвою може мати суттєво різний мінеральний склад, або відрізнятися пористістю, чи іншими характеристиками.

Також окремо стоїть розмірність площі вимірювання, оскільки в межах однієї території можуть проводитися дуже багато вимірювань, а також вона потребує своїх атрибутів тому її включення в таблицю фактів є неможливим. В її межах можуть виділятися певні частини, які матимуть дещо інші характеристики по середній висоті чи іншим характеристикам.

Наступною розмірністю йде Array Number (тип вимірювання) ця назва йде з вхідного файлу формату.dat, Оскільки сам вибір електродної решітки для дослідження питомого опору залежить від багатьох змінних, включно із силою аномального відгуку решітки на тип цілі, на яку проводять дослідження.

Вибір електродної решітки для дослідження індукованої поляризації або питомого опору контролюється багатьма змінними, включаючи силу аномального відгуку решітки на тип цілі, яку шукають, швидкість покриття, рельєф місцевості, потужність і чутливість обладнання і так далі (*Saydam, Duckworth, 1978*).

У своєму дослідженні автори проводили вимірювання горизонтальних профілі IP і уявного питомого опору для решіток Веннера, дипольно-дипольної, дипольно-полюсної, екваторіально-дипольної, квадратно-дипольної і електронної (рухомої поперечної) решіток. Вони проводили оцінювання їх ефективності, де ці методи оцінювалися на основі вимірювань на масштабній моделі над тонкою табличною мішенню. Глибину до вершини (d) і кут нахилу (θ) мішені змінювали, щоб вивчити їхній вплив на реакції IP і питомого опору (*Saydam, Duckworth, 1978*).

Також було виділено стратиграфічну колонку, як окрему розмірність оскільки одні й ті самі породи можуть бути утворені в різні періоди. Як було згадано вище породи, які внесені в таблицю фактів мусять переважати в розрізі, який було проаналізовано на основі методів геоелектрики.

Взагалі, питомий опір геологічних матеріалів має один з найбільших діапазонів серед усіх фізичних властивостей - від $1,6 \times 10^{-8}$ Ω м. для самородного срібла до 10^{16} Ω м. для чистої сірки. Вік породи також є важливим фактором: вулканічна порода четвертинного періоду може мати питомий опір у діапазоні 10-200 м, тоді як питомий опір еквівалентної породи, але докембрійського віку, може бути на порядок більшим. Це є наслідком того, що старіша порода набагато довше піддається вторинному заповненню проміжків мінералізацією та ущільненню, що зменшує пористість і проникність (*Reynolds, 2011, с. 290*).

Наступна розмірність це підземні води. Оскільки чим вище вони до поверхні тим суттєвіше можуть впливати на результати вимірювання. Дану

розмірність варто виділяти через те, що потрібно чітко розуміти мету досліджень це пошук водоносних горизонтів чи порід. І якщо вода стоїть дуже високо до поверхні, то вона суттєво впливати на результати вимірювання.

В рамках своєї роботи автори роботи: «Характеристика водоносної системи з використанням електричної резистивної томографії (ERT) та методів індукованої поляризації (IP)». Проводили, як пошуки так і одразу характеризували водоносні горизонти в Малайзії. Вони застосовували двовимірні (2D) дослідження ERT та IP для отримання важливої інформації для планування розвідки підземних вод та оцінки неглибоких підземних профілів. Розподіл питомого опору в ґрунті може змінюватися залежно від вмісту води та літологічного матеріалу. Зворотне моделювання (інверсія) застосовується для вимірювання справжнього розподілу питомого опору, щоб створити модель на основі видимих даних. На питомий опір ґрунтових матеріалів впливає вміст металів у породі, пористість, вміст глини, проникність і ступінь насиченості пор. В своїй роботі вони прийшли до висновку, що система водоносних горизонтів характеризується перешаруванням піщано-мулистої товщі на досліджуваній території (*N.Akhtar, M.S.Mislan, M.I.Syakir, M.T.Anees, M.S.M. Yusuff, 2021*).

Після розмірності погоди йде розмірність приладу для вимірювання електричного опору. Дану розмірність виділяють оскільки, вольтаж, частота проходження току, а також джерело току, вид постійний чи змінний можуть впливати на виміри. Про це наголошує багато авторів у роботі під назвою «IEEE Guide for Measuring Earth Resistivity, Ground Impedance, and Earth Surface Potentials of a Grounding System». Ця робота була видана в Інститут інженерів електротехніки та електроніки (*Institute of Electrical and Electronics Engineers скорочено IEEE*).

Також варто виділяти дану розмірність оскільки, як можна побачити на рис. 3.6. Прилад ERT має багато складових відповідно їх відмінність між собою, може мати вплив на вимірний опір.



Рис.3.6 Приклад всіх складових для вимірювання опору методом ERT (N.Akhtar, M.S.Mislan, M.I.Syakir, M.T.Anees, M.S.M. Yusuff, 2021)

Далі йде розмірність оператора вимірювального пристрою. Оскільки людський чинник ніхто ніколи не з може, скасувати в тих процесах, де присутня людина. Із цим чинником важко сперечатися. Звісно, можна, сказати, що досвідчений фахівець навряд чи допустить критичних помилок, але, буває, різне. А от про стажера вже важко сказати допустить він помилку чи ні. Головним для даної таблиці є кількість років скільки даний фахівець працює в галузі, є найважливішим оскільки на його базі можна складати думку про репрезентативність виміряних цим оператором даних. Звісно, вплив оператора при дотриманні методів вимірювання буде незначний, але якщо йти шляхом аналізу всіх чинників тоді не потрібно виключати його також.

Далі йде розмірність погоди, оскільки важко чітко визначити, як сильно будуть насичені водою породи на момент проведення вимірювань тому потрібно виділити такий вплив окремо. На основі інформації про погоду

можна спрогнозувати, наскільки сильно породи наситяться водою, чи навпаки вода з них випарується.

У своєму дослідженні під назвою «усунення впливу температури на дані томографії електричного опору, зібрані в частково замерзлому ґрунті: Обмеження та міркування щодо польового застосування» автори продемонстрували, що навіть за ідеальних умов не можна задовільно усунути вплив сезонно замерзлого поверхневого шару з даних, зібраних під час експерименту з моніторингу ERT у часі. Стратегії збору та обробки даних ERT, які часто використовуються в мерзлих ґрунтах, хоча зазвичай дають більш геологічно обґрунтовані інвертовані моделі, не помітно покращують результати температурної корекції, а в деяких випадках призводять до збільшення помилки в даних, скоригованих на температуру. Аналіз чутливості показав, що помилки в даних, скоригованих за температурою, у першу чергу пов'язані з помилкою в оцінній моделі питомого опору, і що ця чутливість до помилки моделі була набагато вищою в частково замерзлому ґрунті, ніж у незамерзлому ґрунті. Зрештою, успішне усунення впливу температури з даних ERT, зібраних у замерзлому ґрунті, вимагає більш точного оцінювання питомого опору ґрунту, ніж та, що досягається на основі сучасних методів збору та обробки даних. (*Herring, Pidlisecky, Cey, 2021*)

Останню розмірність, яку було виділено це агрогрупи (типи ґрунтів). Оскільки електророзвідка, як метод має не велику проникність тому на виміряні значення опору можуть впливати різні характеристики ґрунтів. Зокрема подружжя Позднякових в своїй роботі 2002 року прийшли до висновків, що геофізичні методи не вимірюють окремі заряди в ґрунтах, а окреслюють місця з різною щільністю електричних зарядів. Об'ємна густина рухомих електричних зарядів у ґрунтах вказує на вміст іонів, які нейтралізують заряди на вільній поверхні (*POZDNYAKOV A., POZDNYAKOVA L., 2002*).

Заряд ґрунту визначається іонним обміном, на який впливають хімічні властивості ґрунту, такі як вміст гумусу, насиченість основами, ємність катіонного обміну, мінеральний склад ґрунту та кількість розчинних солей (POZDNYAKOV A., POZDNYAKOVA L., 2002).

Фізичні властивості ґрунту впливають на рухливість електричних зарядів у ґрунті. Об'ємна густина рухомих електричних зарядів пов'язана з багатьма фізичними та хімічними властивостями ґрунту (POZDNYAKOV A., POZDNYAKOVA L., 2002).

Зв'язок між властивостями ґрунту та електричними параметрами не є чітким при вимірюванні декількох властивостей ґрунту одночасно, але загальні експоненціальні залежності були отримані для багатьох властивостей ґрунту (POZDNYAKOV A., POZDNYAKOVA L., 2002).

Й останньою і головною таблицею в цій БД є таблиця фактів. У неї вносяться вся інформація з усіх розмірностей. Через це атрибути в цій таблиці можна розділити на два типи ті, які є унікальними й містяться тільки в ній і ті, які можуть повторюватися в ній кілька разів, оскільки вони існують в інших таблицях.

Також варто дуже уважно розглянути всі атрибути, які мають розмірності, а також типи даних, які були для них обрані й джерела інформації, які будуть використані при заповненні таблиць.

3.3 Заповнення таблиць

Не варто випускати з уваги той факт, що атрибути, якими заповнені розмірності та типи даних для них є теж дуже важливим аспектом розроблення багатовимірної бази даних, як частини системи підтримки прийняття рішень.

Першу розмірність, яку було виділено це площі вона має наступні атрибути: координати полігону містяться в геометричному типі полігон, що записаний у 2D просторі, система координат текстом, номер дозволу на розроблення родовища або на проведення вимірювань записується числом, (за потреби можна встановити текст, якщо дозвіл містить серію) території досліджуваної території в гектарах числом, середня висота на рівнем моря числом. Дана розмірність заповнюється перед вимірювання деякі атрибути на основі документів, а деякі на основі геодезичних вимірювань. Код для створення даної розмірності за допомогою мови запитів SQL та СКБД PostgreSQL показано в додатку А.1. Приклад заповнення теоретичними даними цієї розмірності показано на рис. 3.7.

	ID_area [PK] bigint	Area_Coordinate polygon	System_coor text	Name text	Permit_num bigint	Size_area bigint	Average_height bigint
1	1	((5470437.58,4329394.44),(54703...	CK-63	Poltava_base	245	2000	105
2	2	((5476594.04,4321219.83),(54764...	CK-63	Kremenchuk_base	247	1500	145
3	3	((5478046.006,231491.176),(5478...	УСК-2000	Bugaivka_pond	5678	3000	130
4	4	((5486359.373,224598.458),(5486...	УСК-2000	Pronozivka_basement	1490	8000	155

Рис.3.7 Заповнення таблиці площ.

Через особливість заповнення даних типом полігон в даній розмірності треба привести приклад SQL запиту для наповнення даними.

```
INSERT INTO "public"."DIM_Area" ("Area_Coordinate", "System_coor",
>Name", "Permit_num", "Size_area", "Average_height") VALUES ('((5486359.373
,224598.458),(5486331.603,224106.798))'::polygon,'УСК-2000', 'Pronozivka, 1490,
8000,155);
```

Наступною розмірністю йде Array Number ця назва йде з вхідного файлу формату.dat, який було взято, як основу для створення даної БД. Вона має на меті уточнити, який метод використано для вимірювання електричного опору. Містить такі атрибути: номер методу оскільки у вхідному файлі вони нумеруються не по порядку й повна назва методу записується текстом. Запит

для створення показано в додатку А.2. Приклад заповнення показано на рис. 3.8.

id_Array_num [PK] bigint	Number integer	Description text
1	0	mixed measurements
2	1	Wenner (alpha)
3	2	Pole-pole
4	3	Inline dipole-dipole
5	4	Wenner (beta)
6	5	Wenner (gamma)
7	6	Pole-dipole
8	7	Wenner-Schlumberger
9	8	Equatorial dipole-dipole
10	11	Non-conventional/General array
11	12	Cross-borehole survey (apparent resistivity values)
12	13	Cross-borehole survey (resistancy values)
13	15	Gradient array (only used as sub-array number wit...

Рис. 3.8 Заповнення таблиці тип вимірювання (Array Number)

Далі йде таблиця геохронологічної шкали вона містить інформації про еон, еру, період, епоху вік усе текстом та кількість років у мільйонах числом. За можливості або потреби, звісно, дану розмірність можна доповнити більш детальними атрибутами зокрема, ярус, під ярус і так далі. Дані для даної таблиці не були представлені в якості вхідних для даної роботи, тому ці дані потрібно заповнювати окремо. Вхідні дані для даної розмірності відсутні через це вона наповнюється тією інформацією про породи, які переважають на розрізі, який отримано в рамках дослідження. Запит для створення даної розмірності показано в додатку А.3. Заповнення даної розмірності показано на рис. 3.9.

ID_chronostratigraphic [PK] bigint	Eon text	Era text	Period text	Epoch text	Age text	numerical_age bigint
1	Phanerozoic	Kainozoi	Paleogen	Paleochen	Zelands	62
2	Phanerozoic	Kainozoi	Paleogen	Paleochen	Tanetsky	59
3	Phanerozoic	Kainozoi	Paleogen	Paleochen	Ypres	56
4	Phanerozoic	Kainozoi	Paleogen	Еоцен	Лютетский	48
5	Phanerozoic	Kainozoi	Paleogen	Олигоцен	Бартонский	41
6	Phanerozoic	Kainozoi	Paleogen	Олигоцен	Приабонский	38
7	Phanerozoic	Kainozoi	Paleogen	Олигоцен	Рюпельский	34
8	Phanerozoic	Kainozoi	Paleogen	Олигоцен	Хаттский	28

Рис.3.9 Заповнення геохронологічної таблиці

На всіх вище показаних рисунках, які демонструють таблиці можна побачити типи даних.

Четверта по черзі розмірність підземних вод містить наступні атрибути: номер горизонту від поверхні, оскільки на одній площі в одному місці, може, спостерігатися одна кількість горизонтів в іншому місці менша чи більша їх кількість далі йде глибина початку горизонту, тобто на якій глибині від поверхні він починається потім потужність горизонту всі дані записані в числовому вигляді. Інформація про це відсутня в рамках отриманих вхідних даних через це дана таблиця заповнюється на основі даних буріння скважин, якщо такі проводилися, також на основі карт підземних вод. Реалізація даної розмірності запитом SQL показано в додатку А.4. Заповнення даної показано на рис. 3.10

ID_Ground_water [PK] bigint	number_of_aquifers bigint	Thickness_aquifers bigint	depth_from_surface bigint
1	5	4	10
2	5	6	18
3	5	4	27
4	5	8	40
5	5	3	55
6	3	3	15

Рис.3.10 Заповнення таблиці підземних вод.

Розмірності літологія та літологічний підтип заповнюються на основі даних буріння та виходів порід на денну поверхню в регіоні дослідження. Звісно, методами електричної томографії важко встановити точно через, які породи проходить розріз, але на основі геологічної мапи, яка повинна бути обов'язково створена для досліджуваної території в тому чи іншому масштабі можна встановити, які породи переважають на даному розрізі. Також на основі виходів на денну поверхню порід, варто зазначити, щоб ці виходити були, як умога ближче до розрізу, який досліджується ERT.

Розмірність літологічних підтипів має атрибути назви, максимальної та мінімального значення електричного опору зафіксовані на даній території, а

також середні значення, які переважають для даного підтипу. Зберігаються вони в числом. Також в дану розмірність можна записати літологічний склад текстом та пористість у відсотках числом. Створення даної таблиці за допомогою запитів SQL показано в додатку А.5. Заповнення теоретичними даної цієї таблиці показано на рис. 3.11.

ID_subtype [PK] bigint	Max_electric_resistance bigint	Min_electric_resistance bigint	Average_electric_resistance bigint	Porosity bigint	mineral composition text	Name text
1	100	30	70	40	неоднорідний за складом пісковик,...	пісковик гетерогенний
2	100	30	50	25	пісковик, який містить 50-90 % піро...	пісковик туфогенний
3	200	40	150	7	кварц, глинисті мінерали, цемент	Алевроліт
4	205	70	135	9	Al ₂ O ₃ – 39,56; SiO ₂ – 46,5; H ₂ O – 13...	Глина
5	950	150	542	15	Гравій, гальку та валуни	Пісок валунний
6	950	150	542	13	90% SiO ₂ решта домішки	Пісок кварцовий
7	15	5	8	2	тальк із домішками кварцу, слюд, х...	Тальковий сланець
8	15	4	9	3	Глини, глинистих сланці, філіти	Слюдяний сланець
9	15	6	10	5	гідрослюди, кварц (пилуваті часточ...	Глинисті сланці
10	15	6	10	8	Хлорит з домішками кварцу, тальку...	Хлоритові сланці

Рис. 3.11 Заповнення таблиці Літологічних підтипів

Таблиця Літологія містить наступні атрибути: назву породи (текстом), середній опір (в омах на метр числом), номер підтипу (зовнішній ключ числом) та середню потужність (в метрах числом). Створення даної розмірності за допомогою мови запитів SQL показано в додатку А.6. Заповнення даної розмірності показано на рис. 3.12.

ID_Lithology [PK] bigint	Lithology_type text	Average_thickness bigint	Subtypes_ID bigint	Average_resistivity bigint
1	пісковик	100	1	70
2	Пісок	49	2	642
3	Пісок	32	3	542
4	Сланці	300	7	10
5	Сланці	421	9	11

Рис. 3.12 Заповнення таблиці Літологія

Сьома таблиця даної бази даних для вимірювальних пристроїв містить наступні атрибути: назву приладу текстом, кількість каналів АВ та МН числом, частота операцій числом, кількість режимів числом, максимальне споживання електроенергії числом, стандарт захисту текстом, місткість батареї числом, точність вимірювання числом, максимальний вхідний опір числом в МОhm, тип вимірювання, матеріали дротів числом, матеріали електродів числом, кількість жил в дротах числом. Інформація для даної розмірності не була

представлена в якості вхідної її можна отримати з технічних характеристик приладу. Створення даної таблиці продемонстрована в додатку А.7. Приклад заповнення цієї розмірності показано на рис. 3.13

ID_Device [PK] bigint	Name text	Number_channels bigint	Operating_fre bigint	Number_Oper bigint	Max_Power_c bigint	Protection_st text	Battery_cappacity bigint	Accuracy bigint	Input_impeda bigint	Measuring_ty text	Wires_materi text	Electrod_mati text	Wire_channel bigint
1	ADUK-3/4	3	60	20	7000	IP-64	10000	0	50	ERT	aluminium	metal	1
2	ADMT-1B	3	60	20	6000	IP-64	5200	0	50	ERT	aluminium	metal	2
3	ADMT-6B	2	15	10	5200	IP-64	5200	0	45	ERT	copper	aluminium	10

Рис.3.13 Наповнення таблиці приладів.

Наступна таблиця розмірності це оператор вона має такі атрибути: прізвище, ім'я, назва організації, назва посади всі ці атрибути текстові, кількість років досвіду у вимірюванні електричного опору в роках числом. Дану інформацію можна отримати з організаційних документів, або автор вимірювань їх повинен заповнити власноруч. Створення даної таблиці за допомогою мови запитів SQL показано в додатку А.8. Результати наповнення даної розмірності показано на рис 3.14.

ID_Operators [PK] bigint	Surname text	Name text	Organization text	position text	Year_Experience bigint
1	Столпаков	Андрій	ННІ "Інститут геології" при КНУ Тараса Шевченка	Студент	0
2	Хоменко	Руслан	ННІ "Інститут геології" при КНУ Тараса Шевченка	Викладач	10
3	Віршило	Іван	ННІ "Інститут геології" при КНУ Тараса Шевченка	Викладач	5

Рис.3.14 Заповнення таблиці Оператор.

Наступну розмірність погоди має наступні атрибути: кількість опадів за день, тиждень, місяць до вимірювання в мм., середньодобова температура за день, тиждень, місяць до вимірювання в °С. Дані для наповнення даної інформації можна отримати з гідрометцентру, або мережі інтернет, як історичні дані за певний період та певну територію. Створення цієї розмірності показано в додатку А.9. Наповнення теоретичними даними цієї таблиці показано на рис. 3.15.

ID_weather [PK] bigint	precipitation_day_before bigint	averadge_mm_week bigint	averadge_mm_month bigint	averadge_temp_day_before bigint	averadge_temp_week_before bigint	averadge_temp_month_before bigint
1	5	15	100	8	6	4
2	10	23	145	10	15	14
4	4	50	300	15	15	16
5	29	10	30	10	11	8

Рис.3.15 Заповнення таблиці погодні умови.

Останньою таблицею розмірностей, яку було виділено в рамках багатовимірної бази даних типу ROLAP це агрогрупи ґрунтів. Ця таблиця має наступні атрибути: шифр текстом, повна назва текстом, вміст гумусу % числом, кислотність числом, засоленість числом, пористість числом, питома маса числом, об'ємна маса числом, потужність кожного досліджуваного шару числом, координата шурфа геометричний тип точка. Дані для заповнення цієї розмірності отримуються шляхом викопування шурфа, та лабораторних досліджень всіх отриманих зразків ґрунту. Створення даної розмірності продемонстровано в додатку А.10. Представлені дані на рисунку є теоретичними. Вони продемонстровані на рис. 3.16.

	ID_AgroType [PK] bigint	Encryption text	Full_Name text	Humus_conte bigint	Acidity bigint	Salinization bigint	Porosity bigint	Specific_mas bigint	Volumetric_r bigint	Height bigint	Coordinate_Exmaples_from point
1	5	52Г	Чорноземи типо...	5	10	3	15	31	42	2	(5486191.383,224609.118)
2	6	121г	Лучно-чорнозем...	10	5	10	7	90	10	4	(5486156.153,224612.548)
3	7	13г	Підзолисто-дерн...	14	8	10	7	95	5	3	(5485068.503,225703.608)
4	8	40в	Темно-сірі опідз...	13	11	13	22	98	2	5	(5485068.503,225703.608)
5	9	54е	Чорноземи типо...	20	4	6	8	80	20	7	(5483210.933,226222.298)
6	10	11г	Дерново-підзоли...	13	5	34	14	88	12	3	(5480452.413,224063.058)
7	11	13в	Підзолисто-дерн...	14	6	35	15	89	11	4	(5480452.413,224063.058)

Рис.3.16 Заповнення таблиці Агрогрупи ґрунтів.

Оскільки в даній розмірності також використовується геометричний тип даних варто зазначити приклад запиту для наповнення даної таблиці.

```
INSERT INTO "public"."DIM_AgroType" ("Encryption", "Full_Name",
"Humus_content", "Acidity", "Salinization", "Porosity", "Specific_mass",
"Volumetric_mass", "Height", "Coordinate_Exmaples_from") VALUES ('13в',
'Підзолисто-дернові супіщані ґрунти', 14 ,6, 35, 15, 89, 11, 4,
('(5480452.413,224063.058')::point);
```

Останньою таблицею, яка була створена в рамках даної бази даних це факти, які безпосередньо вимірюються під час проведення вимірювання в польових умовах. Оскільки це таблиця кінцева в неї входять всі дані з інших таблиць через ці атрибути в ній можна поділити на два типи, де перший це дані вимірювання, а другий ID з інших таблиць. До першого типу належать наступні атрибути: фактор «а», фактор «н», виміряне значення електричного

опору, відстань між електродами MN, загальна кількість точок у вхідному файлі, значення IP поляризації чи присутні вони чи ні, Тип «х» розташування для точкових даних використовується для позиціонуванні псевдорозрізу, розташування точки знімання по профілю точкові дані, глибина, на яку вимірюється опір. До другої групи атрибутів даної таблиці належать: ID площі, ID типу вимірювання, ID геохронологічної шкали, ID підземних вод, ID літології, ID приладу, ID оператора, ID погоди, ID Агрогрупи. Створення даної таблиці показано в додатку А.11. Приклад заповнення головної таблиці фактів показано в Додатку Б.

Деякі дані в цій таблиці є теоретичним, як наслідок з інших таблиць. Також на даному рисунку не розміщені всі атрибути через їх велику кількість.

Таким чином була побудована багатовимірна база даних, як складова ROLAP системи у рамках системи керування базами даних PostgreSQL, для даних електрорезистивної томографії. Було виділено 10 розмірностей. Що дає можливість, для їх аналізу окремо не залежно один від одного, що у свою чергу, може, дати нові гіпотези та нові знання про середовище.

ВИСНОВКИ

В рамках даної роботи всі поставлені задачі були виконані. Та були отримані наступні про те, що певний перелік факторів впливає на вимірний електричний опір, який отримують різними методами. Зокрема: спосіб вимірювання та вимірювальна сітка, підземні води, літологічні типи та підтипи, вік порід абсолютний та відносний, вимірювальний пристрій, оператор, погодні умови до вимірювань та ґрунти.

Ці всі фактори були виділені на основі аналізу методів електророзвідки, які в свою чергу маю розв'язувати пряма та обернену задачу. Які не взагалі не мають аналітичного розв'язку через це дослідники намагаються різними методами знайти закономірності в опорі та породами, що і було метою даної роботи.

На основі тих факторів, які були виділені, які такі, що мають вплив на вимірний електричний опір. Було побудовано сховище даних (data warehouse) типу ROLAP – реляційна онлайн-аналітична обробка. І на базі реляційної системи керування базами даних PostgreSQL була побудована багатовимірна база даних типу сніжинка. Де було створено 11 таблиць з них 9 головних розмірностей і одна ієрархічна від іншої.

Розроблене сховище даних можна використати для комплексного аналізу впливу різних факторів на вимірний опір. Великим недоліком даної роботи, є відсутність реальних даних, якими можна було б заповнити всі розмірності та проаналізувати всі переваги застосування даного підходу для даних електрорезистивної томографії.

На мою думку, якщо реалізувати в повній мірі розроблену схему багатовимірної бази даних у поєднанні з великими масивами даних, як вимірних (тобто заповнення таблиці фактів), так і заповнення всіх

розмірностей разом з використанням алгоритмів глибокого навчання можна отримати нові знання та закономірності виміряного опору.

Отже, можна зробити висновок про те, що використання систем прийняття рішень і БД має вирішальне значення для успіху в будь-якій сфері, у тому числі й у геології. Багатовимірні бази даних стають усе більш важливими для систем прийняття рішень, оскільки вони можуть обробляти складні структури даних і дозволяють швидше та ефективніше робити запити та аналізувати великі набори даних. У геології використання багатовимірних БД, може, дає змогу геологам краще зрозуміти структуру і властивості земних надр і приймати обґрунтовані рішення щодо розвідки і видобутку корисних копалин.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ:

КНИГИ
Один автор або декілька авторів
Alter, S. (2013). Decision support systems: Current practice and continuing challenges. In Handbook on decision support systems 1 (pp. 1-33). Springer US. Addison-Wesley; First edition (January 1, 1980) ISBN-10: 0201001934. ISBN-13 : 978-0201001938
Galliers, R. D., & Leidner, D. E. (2009). Strategic information management: challenges and strategies in managing information systems. Routledge. ISBN-10 0415996473. ISBN-13: 978 0415996471. 584 стр.
Hammond, J. S., Keeney, R. L., & Raiffa, H. (1999). Smart choices: A practical guide to making better decisions. Harvard Business Press
Janis, I. L., & Mann, L. (1977). Decision making: A psychological analysis of conflict, choice, and commitment. Free Press.
Meredith, J. R., & Mantel Jr, S. J. (2012). Project management: a managerial approach. https://www.academia.edu/8973287/Project Management A Management Approach Jack R Meredith and Samuel J Mantel Wiley . ISBN-13 978-0-470-22621-6. Copyright © 2009 John Wiley & Sons, Inc.
Paul C. Nutt , David C. Wilson.(2010)Handbook of Decision Making. ISBN: 978-1-444-32312-2. Wiley-Blackwell 720 Pages
Power, Daniel J., "Decision Support Systems: Concepts and Resources for Managers" (2002). Faculty Book Gallery. 67.
Reynolds, J. M. (2011). An Introduction to Applied and Environmental Geophysics (2-ге вид.). Willey-BlackWell A John Wiley & Sons, Ltd., Publication. (Оригінал опубліковано 1997 р.)
Salley, C., & Codd, E.F. (1998). Providing OLAP to User-Analysts: An IT Mandate.,

<p>Sharda, R., Delen, D., & Turban, E. (2014). Decision support and business intelligence systems. Upper Saddle River, NJ: Pearson Education. https://www.academia.edu/download/62386890/business_intelligence20200316-50198-128rsmo.pdf</p>
<p>Simon, H. A. (1960). The new science of management decision. Harper and Row. Volume 3 of Ford distinguished lectures. Видавництво: John Wiley & Sons.</p>
<p>Zhou, B. (2019). Electrical Resistivity Tomography: A Subsurface-Imaging Technique. IntechOpen. doi: 10.5772/intechopen.81511</p>
<p>Частина книги (незалежно від кількості авторів)</p>
<p>Pal, K. (2019). Quality Assurance Issues for Big Data Applications in Supply Chain Management. <i>У Predictive Intelligence Using Big Data and the Internet of Things</i> (с. 51–76). IGI Global. https://doi.org/10.4018/978-1-5225-6210-8.ch003</p>
<p>ПЕРІОДИЧНІ, ПРОДОВЖУВАНІ ВИДАННЯ, МАТЕРІАЛИ КОНФЕРЕНЦІЙ</p>
<p>Стаття в журналі (до шести авторів)</p>
<p>Akhtar, N., Mislán, M. S., Syakir, M. I., Anees, M. T., & Yusuff, M. S. M. (2021). Characterization of aquifer system using electrical resistivity tomography (ERT) and induced polarisation (IP) techniques. <i>IOP Conference Series: Earth and Environmental Science</i>, 880(1), 012025. https://doi.org/10.1088/1755-1315/880/1/012025</p>
<p>Alyae, S., Suter, E., Bratvold, R. B., Hong, A., Luo, X., & Fossum, K. (2019). A decision support system for multi-target geosteering. <i>Journal of Petroleum Science and Engineering</i>, 183, 106381. doi:10.1016/j.petrol.2019.106381</p>
<p>Dahlin, T., & Zhou, B. (2004). A numerical comparison of 2D resistivity imaging with 10 electrode arrays. <i>Geophysical Prospecting</i>, 52(5), 379–398. https://doi.org/10.1111/j.1365-2478.2004.00423.x</p>
<p>Ettore Cardarelli, Giorgio De Donno. Multidimensional electrical resistivity survey for bedrock detection at the Rieti Plain (Central Italy). <i>Journal of Applied</i></p>

<p>Geophysics Volume 141, June 2017, Pages 77-87. https://doi.org/10.1016/j.jappgeo.2017.04.012</p>
<p>Florian Wellmann & Guillaume Caumon. 3-D Structural Geological Models: Concepts, Methods, and Uncertainties. <i>Advances in Geophysics</i> (2018), Vol. 59, Chap. 1, pp. 1–121, Elsevier. https://doi.org/10.1016/bs.agph.2018.09.001.</p>
<p>GHEORGHE, D., & ARMAS, I. (2015). GIS BASED DECISION SUPPORT SYSTEM FOR SEISMIC RISK IN BUCHAREST. CASE STUDY – THE HISTORICAL CENTRE. <i>Journal of Engineering Studies and Research</i>, Volume 21(3), 35–42. https://doi.org/10.29081/jesr.v21i3.16</p>
<p>Gorla, M., Simonetti, R., & Righetti, C. (2016). Basin-scale hydrogeological, geophysical, geochemical and isotopic characterization: an essential tool for building a Decision Support System for the sustainable management of alluvial aquifer systems within the provinces of Milan and Monza-Brianza. <i>Acque Sotterranee - Italian Journal of Groundwater</i>, 5(3). https://doi.org/10.7343/as-2016-213</p>
<p>He, Y., Tian, D., Wang, H., Yao, L., Yu, M., & Chen, P. (2019). A universal and multi-dimensional model for analytical data on geological samples. <i>Geoscientific Instrumentation, Methods and Data Systems</i>, 8(2), 277–284. https://doi.org/10.5194/gi-8-277-2019</p>
<p>Herring, T., Pidlisecky, A., & Cey, E. (2021). Removing the effects of temperature on electrical resistivity tomography data collected in partially frozen ground: Limitations and considerations for field applications. <i>Vadose Zone Journal</i>. https://doi.org/10.1002/vzj2.20146</p>
<p>James S. Dyer, Richard N. Lund, John B. Larsen, V. Kumar and Robert P. Leone. A Decision Support System for Prioritizing Oil and Gas Exploration Activities. <i>Operations Research</i> Vol. 38, No. 3 (May - Jun., 1990), pp. 386-396 (11 pages) Published By: INFORMS https://www.jstor.org/stable/171353</p>

Kumar, V., & Reinartz, W. (2012). Creating enduring customer value. *Journal of Marketing*, 80(6), 1-18. <https://doi.org/10.1509/jm.15.041>

Morfonios, K., Konakas, S., Ioannidis, Y., & Kotsis, N. (2007). ROLAP implementations of the data cube. *ACM Computing Surveys*, 39(4), 12–es. doi:10.1145/1287620.1287623

Pedersen, T. B., & Jensen, C. S. (2001). Multidimensional database technology. *Computer*, 34(12), 40–46. doi:10.1109/2.970558

Saydam, A. S., & Duckworth, K. (1978). Comparison of some electrode arrays for their IP and apparent resistivity responses over a sheet like target. *Geoexploration*, 16(4), 267–289. [https://doi.org/10.1016/0016-7142\(78\)90016-9](https://doi.org/10.1016/0016-7142(78)90016-9)

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124-1131.

Стаття в журналі (більше шести авторів)

Chambers, J. E., Gunn, D. A., Wilkinson, P. B., Meldrum, P. I., Haslam, E., Holyoake, S., Kirkham, M., Kuras, O., Merritt, A., & Wragg, J. (2012). 4D electrical resistivity tomography monitoring of soil moisture dynamics in an operational railway embankment. *Near Surface Geophysics*, 12(1), 61–72. <https://doi.org/10.3997/1873-0604.2013002>

El Bazi, N.; Mabrouki, M.; Laayati, O.; Ouhabi, N.; El Hadraoui, H.; Hammouch, F.-E.; Chebak, A. Generic Multi-Layered Digital-Twin-Framework-Enabled Asset Lifecycle Management for the Sustainable Mining Industry. *Sustainability* 2023, 15, 4. <https://doi.org/10.3390/su15043470>

Singh, U., & Sharma, P. K. (2022). Study on geometric factor and sensitivity of subsurface for different electrical resistivity Tomography Arrays. *Arabian Journal of Geosciences*, 15(7). <https://doi.org/10.1007/s12517-022-09844-3>

Тези та матеріали конференцій

Aliki Kokkala, V. Marinos Engineering geological and geotechnical big data management for planning and protecting smart cities: the pilot case of the city of Thessaloniki. October 2022 Conference: 16th International Congress of the Geological Society of Greece, Our Earth, Our Home At: Patras, Greece

Günther, T., & Rucker, C. (2012). Electrical Resistivity Tomography (ERT) in geophysical applications - state of the art and future challenges. Conference: Schlumberger Symposium - 100 years of electrical imaging, Paris

POZDNYAKOV, A., & POZDNYAKOVA, L. (2002). Electrical fields and soil properties. *У Conference: 17 World Congress of Soil Science* (с. 1558). [https://www.researchgate.net/publication/227339828 Electrical fields and soil properties](https://www.researchgate.net/publication/227339828_Electrical_fields_and_soil_properties)

Shastri L Nimmagadda and Amit Rudra. Data sources and requirement analysis for multidimensional database modeling – an Australian Resources Industry scenario. January 2004. Conference: proceedings in the 7th international conference on IT. At: Hyderabad, India

ЕЛЕКТРОННІ РЕСУРСИ

Електронні ресурси (сайти)

What is CRM? (б. д.). Офіційний сайт Salesforce. <https://www.salesforce.com/crm/what-is-crm>

The leader in collaboration & customer experience | оф. Сайт Webex. (б. д.). Webex. <https://www.webex.com>

Oracle Hyperion Planning | Enterprise Performance Management | оф. Сайт Oracle СНГ. (б. д.). Oracle | Cloud Applications and Cloud Platform. <https://www.oracle.com/cis/performance-management/hyperion-planning/>

Rouse, M. (2017, 18 січня). <i>Multidimensional Database</i> . Techopedia. https://www.techopedia.com/definition/14710/multidimensional-database-mdb
Учасники проєктів Вікімедіа. (2010, 31 січня). <i>OLAP-куб</i> — Вікіпедія. Вікіпедія. https://uk.wikipedia.org/wiki/OLAP-куб
Dantoni, J. (2022, 12 травня). <i>MOLAP Defined: Architecture, Advantages & Disadvantages</i> . Oracle NetSuite. https://www.netsuite.com/portal/resource/articles/data-warehouse/multidimensional-online-analytical-processing-molap.shtml
Mbaabu, O. (2021, 24 січня). <i>MOLAP vs ROLAP vs HOLAP in Online Analytical Processing (OLAP)</i> . Engineering Education (EngEd) Program Section. https://www.section.io/engineering-education/molap-vs-rolap-vs-holap/
TechTarget Contributor (2021, 21 травня). <i>What is a star schema and how does it work?</i> Data Management. https://www.techtarget.com/searchdatamanagement/definition/star-schema
<i>Snowflake Schema in Data Warehouse Model - GeeksforGeeks</i> . (2023, 9 березня). GeeksforGeeks. https://www.geeksforgeeks.org/snowflake-schema-in-data-warehouse-model/
Pedamkar, P. (б. д.). <i>HOLAP Architecture of HOLAP with Advantages & Disadvantages</i> . EDUCBA. https://www.educba.com/holap/
Loke MH (2004) Tutorial: 2-D and 3-D electrical imaging surveys. http://www.geotomosoft.com/ .
8.8. <i>Geometric Types</i> . (б. д.). PostgreSQL Documentation. https://www.postgresql.org/docs/current/datatype-geometric.html

What is VES, ERT, IP and SP Electrical Resistivity ? - Seis Tech. (б. д.). Seis Tech. <https://www.seis-tech.com/faq/what-is-ves-ert-ip-sp-electrical-resistivity/>

РІЗНЕ

Стандарти та нормативні документи

Benito Abril Cogque IEEE Guide for Measuring Earth Resistivity, Ground Impedance, and Earth Surface Potentials of a Grounding System Sponsored by the Substations Committee IEEE Power and Energy Society. IEEE3 Park Avenue New York, NY 10016-5997 USA28 December 2012 IEEE Std 81™-2012 (Revision of IEEE Std 81-1983)

Додаток А

Додаток А.1

SQL код для створення таблиці Площі.

```
CREATE TABLE IF NOT EXISTS public."DIM_Area"
("ID_area"      bigint      NOT      NULL      DEFAULT      nextval
"Area_Coordinate_ID_area_seq"::regclass),
  "Area_Coordinate" polygon,
  "System_coor" text COLLATE pg_catalog."default",
  "Name" text COLLATE pg_catalog."default",
  "Permit_num" bigint,
  "Size_area" bigint,
  "Average_height" bigint,
  CONSTRAINT "PM_Area" PRIMARY KEY ("ID_area"))
TABLESPACE pg_default;
ALTER TABLE IF EXISTS public."DIM_Area"
  OWNER to postgres;
```

Додаток А.2

Sql запит для створення розмірності тип сітки (Array Number):

```
CREATE TABLE IF NOT EXISTS public."DIM_Array_number"
(
  "id_Array_number"      bigint      NOT      NULL      DEFAULT
  nextval("Array_number_id_Array_number_seq"::regclass),
  "Number" integer NOT NULL,
  "Description" text COLLATE pg_catalog."default" NOT NULL,
  CONSTRAINT "PM_Array_number" PRIMARY KEY ("id_Array_number"))
TABLESPACE pg_default;
ALTER TABLE IF EXISTS public."DIM_Array_number"
  OWNER to postgres;
```

Додаток А.3

SQL Запит для створення таблиця геохронологічної шкали.

```

CREATE TABLE IF NOT EXISTS public."DIM_Chronostratigraphic"
( "ID_chronostratigraphic" bigint NOT NULL,
  "Eon" text COLLATE pg_catalog."default",
  "Era" text COLLATE pg_catalog."default",
  "Period" text COLLATE pg_catalog."default",
  "Epoch" text COLLATE pg_catalog."default",
  "Age" text COLLATE pg_catalog."default",
  "numerical_age_(Ma)" bigint,
  CONSTRAINT "PM_Chronostratigraphic" PRIMARY KEY
("ID_chronostratigraphic"))
TABLESPACE pg_default;
ALTER TABLE IF EXISTS public."DIM_Chronostratigraphic"
OWNER to postgres;

```

Додаток А.4

```

CREATE TABLE IF NOT EXISTS public."DIM_Ground_water"
( "ID_Ground_water" bigint NOT NULL,
  number_of_aquifers bigint,
  "Thickness_aquifers" bigint,
  depth_from_surface bigint,
  CONSTRAINT "PM_ground_water" PRIMARY KEY ("ID_Ground_water"))
TABLESPACE pg_default;
ALTER TABLE IF EXISTS public."DIM_Ground_water"
OWNER to postgres;

```

Додаток А.5

```

CREATE TABLE IF NOT EXISTS public."DIM_Lithology_subtype "
( "ID_subtype" bigint NOT NULL,
  "Name" bigint,
  "Max_electric_resistance" bigint,
  "Min_electric_resistance" bigint,
  "Avareadge_electric_resistance" bigint,
  "Porosity" bigint,
  "mineral composition" text COLLATE pg_catalog."default",
  CONSTRAINT "PM_subtype" PRIMARY KEY ("ID_subtype"))
TABLESPACE pg_default;
ALTER TABLE IF EXISTS public."DIM_Lithology_subtype "
OWNER to postgres;

```

Додаток А.6

```

CREATE TABLE IF NOT EXISTS public."DIM_Lithology"
( "ID_Lithology" bigint NOT NULL,
  "Lithology_type" text COLLATE pg_catalog."default",
  "Average_thickness " bigint,
  "Subtypes_names" bigint,
  "Average_resistivity" bigint,
  CONSTRAINT "PM_Lithlogy" PRIMARY KEY ("ID_Lithology"),
  CONSTRAINT "FK_subtypes" FOREIGN KEY ("Subtypes_names")
    REFERENCES public."DIM_Lithology_subtype " ("ID_subtype") MATCH SIMPLE
    ON UPDATE NO ACTION
    ON DELETE NO ACTION)
TABLESPACE pg_default;
ALTER TABLE IF EXISTS public."DIM_Lithology"
  OWNER to postgres;

```

Додаток А.7

```

CREATE TABLE IF NOT EXISTS public."DIM_Measuring_device"
( "ID_Device" bigint NOT NULL,
  "Name" text COLLATE pg_catalog."default",
  "Number of channels (MN and AB)" bigint,
  "Operating_frequencies" bigint,
  "Number_Operation_modes" bigint,
  "Max_Power_consumption" bigint,
  "Protection_standard" text COLLATE pg_catalog."default",
  "Battery_cappacity" bigint,
  "Accuracy" bigint,
  "Input_impedance_MOhm" bigint,
  "Measuring_type" text COLLATE pg_catalog."default",
  "Wires_material" text COLLATE pg_catalog."default",
  "Electrod_material" text COLLATE pg_catalog."default",
  "Wire_channels" bigint,
  CONSTRAINT "PM_Device" PRIMARY KEY ("ID_Device"))
TABLESPACE pg_default;
ALTER TABLE IF EXISTS public."DIM_Measuring_device"
  OWNER to postgres;

```

Додаток А.8

```
CREATE TABLE IF NOT EXISTS public."DIM_Operators"
( "ID_Operators" bigint NOT NULL,
  "Surname" text COLLATE pg_catalog."default",
  "Name" text COLLATE pg_catalog."default",
  "Organization" text COLLATE pg_catalog."default",
  "position" text COLLATE pg_catalog."default",
  "Year_Experience" bigint,
  CONSTRAINT "PM_Operators" PRIMARY KEY ("ID_Operators"))
TABLESPACE pg_default;
ALTER TABLE IF EXISTS public."DIM_Operators"
OWNER to postgres;
```

Додаток А.9

```
CREATE TABLE IF NOT EXISTS public."DIM_Weather"
( "ID_weather" bigint NOT NULL,
  precipitation_day_before bigint,
  averadge_mm_week bigint,
  averadge_mm_month bigint,
  averadge_temp_day_before bigint,
  averadge_temp_week_before bigint,
  averadge_temp_month_before bigint,
  CONSTRAINT "PM_Wether" PRIMARY KEY ("ID_weather"))
TABLESPACE pg_default;
ALTER TABLE IF EXISTS public."DIM_Weather"
OWNER to postgres;
```

Додаток А.10

```
CREATE TABLE IF NOT EXISTS public."DIM_AgroType"
( "ID_AgroType" bigint NOT NULL,
  "Encryption" text COLLATE pg_catalog."default",
  "Full_Name" text COLLATE pg_catalog."default",
  "Humus_content" bigint,
  "Acidity" bigint,
  "Salinization" bigint,
  "Porosity" bigint,
```

```

"Specific_mass" bigint,
"Volumetric_mass" bigint,
"Height" bigint,
"Coordinate_Exmaples_from" point,
CONSTRAINT "PM_AgroType" PRIMARY KEY ("ID_AgroType"))
TABLESPACE pg_default;
ALTER TABLE IF EXISTS public."DIM_AgroType"
OWNER to postgres;

```

Додаток А.11

```

CREATE TABLE IF NOT EXISTS public."Facts_Measuring_data"
(
    "ID_Measuring"    bigint    NOT    NULL    DEFAULT
nextval("Measuring_data_ID_Measuring_seq"::regclass),
    "A_range"    bigint,
    "N_factor"    bigint,
    measuting_resistance    bigint,
    "ID_Area"    smallint    NOT    NULL,
    "Electrod_range"    bigint,
    "ID_Array_numb"    bigint,
    "Total_number"    bigint,
    "IP_polar"    bigint,
    "Type_X_location"    bigint[],
    "Location"    point,
    "Height"    bigint,
    id_device    bigint,
    "ID_chronostratigraphic"    bigint,
    "ID_operators"    bigint,
    "ID_weather"    bigint,
    "Measuring_depth"    bigint,
    "ID_Lithology"    bigint,
    "ID_Ground_water"    bigint,
    "ID_AgroType"    bigint,
    CONSTRAINT "PM_Measuring"    PRIMARY    KEY    ("ID_Measuring"),
    CONSTRAINT    fk_agrotype    FOREIGN    KEY    ("ID_AgroType")
REFERENCES    public."DIM_AgroType"    ("ID_AgroType")    MATCH    SIMPLE
ON    UPDATE    NO    ACTION
ON    DELETE    NO    ACTION
NOT    VALID,
    CONSTRAINT    fk_area    FOREIGN    KEY    ("ID_Area")

```

```

REFERENCES public."DIM_Area" ("ID_area") MATCH SIMPLE
ON UPDATE NO ACTION
ON DELETE NO ACTION
NOT VALID,
CONSTRAINT fk_array FOREIGN KEY ("ID_Array_num")
REFERENCES public."DIM_Array_number" ("id_Array_number") MATCH
SIMPLE
ON UPDATE NO ACTION
ON DELETE NO ACTION
NOT VALID,
CONSTRAINT fk_chronostratigraphic FOREIGN KEY ("ID_chronostratigraphic")
REFERENCES public."DIM_Chronostratigraphic" ("ID_chronostratigraphic")
MATCH SIMPLE
ON UPDATE NO ACTION
ON DELETE NO ACTION
NOT VALID,
CONSTRAINT fk_device FOREIGN KEY (id_device)
REFERENCES public."DIM_Measuring_device" ("ID_Device") MATCH SIMPLE
ON UPDATE NO ACTION
ON DELETE NO ACTION
NOT VALID,
CONSTRAINT fk_ground_water FOREIGN KEY ("ID_Ground_water")
REFERENCES public."DIM_Ground_water" ("ID_Ground_water") MATCH
SIMPLE
ON UPDATE NO ACTION
ON DELETE NO ACTION
NOT VALID,
CONSTRAINT fk_lithology FOREIGN KEY ("ID_Lithology")
REFERENCES public."DIM_Lithology" ("ID_Lithology") MATCH SIMPLE
ON UPDATE NO ACTION
ON DELETE NO ACTION
NOT VALID,
CONSTRAINT fk_operators FOREIGN KEY ("ID_operators")
REFERENCES public."DIM_Operators" ("ID_Operators") MATCH SIMPLE
ON UPDATE NO ACTION
ON DELETE NO ACTION
NOT VALID,
CONSTRAINT fk_weather FOREIGN KEY ("ID_weather")
REFERENCES public."DIM_Weather" ("ID_weather") MATCH SIMPLE
ON UPDATE NO ACTION
ON DELETE NO ACTION
NOT VALID)

```

```
TABLESPACE pg_default;  
ALTER TABLE IF EXISTS public."Facts_Measuring_data"  
OWNER to postgres;
```

Додаток Б

Приклад заповнення таблиці фактів

ID_Measuring [PK] bigint	A_range bigint	N_factor bigint	measuting_resis bigint	ID_Area smallint	Electrod_rang bigint	ID_Array_num bigint	Total_number bigint	IP_polar bigint	Type_X_locat bigint[]	Location point	Height bigint	id_device bigint	ID_chronostr bigint	ID_operators bigint	ID_weather bigint	Measuring_d bigint	ID_Lithology bigint	ID_Ground_w bigint	ID_AgroType bigint
3	1	1	20	1	15	8	1219	0	{1}	(5470930...	100	1	1	2	1	90	1	1	5
4	1	1	28	1	15	8	1219	0	{1}	(5470930...	100	1	1	2	1	90	1	1	5
5	1	1	90	1	15	8	1219	0	{1}	(5470930...	100	1	1	2	1	90	1	1	5
6	1	1	129	1	15	8	1219	0	{1}	(5470930...	100	1	1	2	1	90	1	1	5
7	1	1	178	1	15	8	1219	0	{1}	(5470930...	100	1	1	2	1	90	1	1	5
8	1	1	264	1	15	8	1219	0	{1}	(5468306...	105	1	1	2	1	90	2	2	7
9	1	1	231	1	15	8	1219	0	{1}	(5468306...	105	1	1	2	1	90	2	2	7
10	1	1	313	1	15	8	1219	0	{1}	(5468306...	105	1	1	2	1	90	2	2	7
11	1	1	292	1	15	8	1219	0	{1}	(5468306...	105	1	1	2	1	90	2	2	7
13	1	1	202	1	30	8	1219	0	{1}	(5468011...	105	2	2	3	4	80	4	3	9
14	1	1	291	1	30	8	1219	0	{1}	(5468011...	105	2	2	3	4	80	4	3	9
15	1	1	153	1	30	8	1219	0	{1}	(5468011...	105	2	2	3	4	80	4	3	9
16	1	1	134	1	30	8	1219	0	{1}	(5468011...	105	2	2	3	4	80	4	3	9