

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

КВАЛІФІКАЦІЙНА РОБОТА МАГІСТРА
Спеціальність 103 – Науки про Землю
Освітня програма: «Геофізика»

ТЕМА: «Цифрові близнюки підземних сховищ газу: перевірка гіпотези подібності геофізичних характеристик на прикладі Полтавського сховища газу»

Виконав	студент 2-го курсу магістратури кафедри геофізика Назарук Іван Олександрович
Науковий керівник	Професор, Доктор геологічних наук, Орлюк Михайло Іванович
Консультант	Доцент кафедри геофізики, доктор геологічних наук, Шабатура Олександр Вікторович

Робота рекомендується до захисту (протокол № 13 засідання кафедри геофізики
від 15-05-2026)

Завідувач кафедри	Доцент кафедри геофізики, кандидат геологічних наук, Онищук Віктор Іванович
-------------------	---

Київ – 2026

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ ...	3
ВСТУП	4
РОЗДІЛ 1.ГЕОЛОГО ГЕОФІЗИЧНА ХАРАКТЕРИСТИКА ПОЛТАВСЬКОГО ПСГ	7
1.1 Географічне положення та історія створення об'єкта.....	7
1.2 Геологічна будова та літолого-петрофізична характеристика району	13
1.3 Аналіз попередніх геолого-геофізичних досліджень та стан вивченості теми	27
РОЗДІЛ 2. ТЕОРЕТИЧНІ ОСНОВИ СТВОРЕННЯ ЦИФРОВИХ БЛИЗНЮКІВ У ГЕОФІЗИЧНИХ ДОСЛІДЖЕННЯХ.....	31
2.1 Поняття цифрового близнюка та його роль у нафтогазовій галузі.....	31
2.2. Основні етапи та програмні засоби побудови цифрових моделей	40
2.3. Світовий досвід використання цифрових близнюків підземних сховищ газу	46
РОЗДІЛ 3. МЕТОДОЛОГІЯ НЕПАРАМЕТРИЧНОЇ КЛАСИФІКАЦІЇ ПРИ ПОБУДОВІ МОДЕЛІ СХОВИЩА	50
3.1 Теоретичні основи кластер-аналізу та самоорганізовуваних мереж Кохонена	50
3.2 Алгоритм підготовки та структурування багатопараметричних масивів даних для ПСГ	53
3.3. Процедури крос-перевірки та верифікації прогностичної здатності моделі	59
РОЗДІЛ 4. ПРАКТИЧНА РЕАЛІЗАЦІЯ ЦИФРОВОГО БЛИЗНЮКА ТА ПЕРЕВІРКА ГІПОТЕЗИ ПОДІБНОСТІ ГЕОФІЗИЧНИХ ХАРАКТЕРИСТИК	65
4.1 Формування та підготовка геолого-геофізичної бази даних для моделювання.....	65
4.2. Побудова моделі непараметричної класифікації ознак колекторів	70
4.3 Аналіз результатів непараметричної класифікації та верифікація моделі на аналогах Перу	82
ВИСНОВКИ	93
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	95
ДОДАТКИ.....	105

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

ПСГ	Підземне сховище газу
ДДз	Дніпровсько-Донецька западина
ГДС	Геофізичні дослідження свердловин
SOM	Self-Organizing Maps (самоорганізовані карти Кохонена)
ML	Machine Learning (машинне навчання)
GR	Gamma Ray (гамма-каротаж)
RHOV	Bulk Density (параметр густинного каротажу)
AC	Acoustic Log (акустичний каротаж)
RES	Resistivity Log (каротаж питомого електричного опору)
LAS	Log ASCII Standard (стандартний формат файлів геофізичного каротажу)
TIFF	Tagged Image File Format (формат графічних файлів)
API	одиниця вимірювання природної радіоактивності порід
Пласт-колектор	проникний геологічний пласт, здатний накопичувати та фільтрувати флюїди
Цифровий близнюк	цифрова динамічна модель об'єкта, що відображає його фізичний стан та поведінку
Непараметрична класифікація	метод класифікації даних без припущення про закон їх статистичного розподілу

ВСТУП

Актуальність теми: Сьогодні питання енергетичної безпеки України набуває особливого значення, особливо в умовах нестабільності світового енергетичного ринку та змін у системах постачання природного газу. Одну з ключових ролей у забезпеченні стабільної роботи газотранспортної системи держави відіграють підземні сховища газу (ПСГ), які дозволяють компенсувати сезонні коливання споживання та підтримувати необхідний рівень газових резервів.

У зв'язку з цим дедалі більшої актуальності набувають питання моніторингу, прогнозування та оптимізації режимів роботи ПСГ. Одним із перспективних напрямів у нафтогазовій галузі є використання цифрових близнюків — динамічних цифрових моделей, які оновлюються на основі геолого-геофізичних і промислових даних та відображають поточний стан об'єкта під час його експлуатації.

Для підземних сховищ газу такий підхід має особливе значення, оскільки ПСГ працюють у циклічному режимі закачування та відбору газу, що супроводжується змінами пластового тиску та гідродинамічного стану колектора. Водночас створення цифрових близнюків ускладнюється великою кількістю різномірних геолого-геофізичних даних та їхньою неоднорідністю.

Саме тому сьогодні все активніше застосовуються методи машинного навчання, кластер-аналізу та самоорганізовані мережі Кохонена, які дозволяють виявляти закономірності у великих масивах даних і визначати ступінь подібності між пластами-колекторами. У межах даної роботи саме ці методи використовуються для перевірки гіпотези подібності геофізичних характеристик на прикладі Полтавського ПСГ.

Мета: полягає у дослідженні можливостей створення цифрового близнюка підземного сховища газу на основі застосування непараметричної кластеризації багатовимірних геофізичних ознак та перевірки гіпотези їхньої подібності на прикладі Полтавського ПСГ.

Для досягнення мети сформульовані такі завдання:

1. Вивчити географічне положення, геологічну будову та літолого-петрофізичні особливості району розташування Полтавського ПСГ.
2. Проаналізувати результати попередніх геолого-геофізичних досліджень та оцінити ступінь вивченості об'єкта.
3. Дослідити теоретичні засади створення цифрових близнюків у нафтогазовій галузі.
4. Проаналізувати можливості застосування математично-статистичних алгоритмів для роботи з багатовимірними геолого-геофізичними даними.
5. Виконати непараметричну кластеризацію геолого-геофізичних даних та визначити найбільш інформативні критерії подібності пластів.
6. Провести верифікацію отриманої моделі та оцінити можливість використання аналогів при побудові цифрового близнюка ПСГ.

Об'єкт дослідження: геологічна будова та геофізичні властивості Полтавського підземного сховища газу та подібних до нього нафтогазових об'єктів Перу.

Предмет дослідження: методи інтеграції багатовимірних геолого-геофізичних даних і застосування непараметричної класифікації при створенні цифрових близнюків.

Методи дослідження. У роботі використано комплексний підхід, який включає методи геолого-геофізичного аналізу, статистичної обробки даних, цифрового моделювання та машинного навчання. Основний акцент зроблено на застосуванні алгоритму кластерного аналізу та автоматизованих нейромереж для побудови багатовимірної кластеризації та аналізу подібності геофізичних характеристик

Особистий внесок автора. Зібрано і систематизовано фактичний матеріал щодо Полтавського ПСГ та нафтогазових родовищ Перу. Обґрунтовано доцільність

застосування методів машинного навчання та непараметричної класифікації для аналізу багатовимірних геолого-геофізичних даних. Виконано обробку даних, проведено кластер-аналіз і побудовано самоорганізовані мережі Кохонена, визначено основні показники подібності між досліджуваними об'єктами та оцінено ефективність створеної прогностичної моделі.

Практичне значення роботи. полягає у можливості використання запропонованого підходу для автоматизації аналізу геолого-геофізичних даних, пошуку аналогів пластів-колекторів та оптимізації процесу створення цифрових близнюків підземних сховищ газу. Отримані результати можуть бути використані під час прогнозування поведінки колекторів у процесі експлуатації ПСГ та для підвищення ефективності управління режимами закачування і відбору газу.

Апробація результатів роботи. Основні результати дослідження були апробовані на XVI Всеукраїнській конференції-школі «Сучасні проблеми наук про Землю» (м. Київ, 8–9 квітня 2026 р.), де було представлено доповідь на тему: *«Непараметрична класифікація багатовимірних геологічних ознак для встановлення генетичних особливостей гранітоїдів Українського щита»*.

У межах доповіді було висвітлено підходи до застосування алгоритму CART для аналізу багатовимірних геолого-геофізичних даних, крос-перевірки моделей та інтерпретації результатів непараметричної класифікації.

РОЗДІЛ 1.ГЕОЛОГО ГЕОФІЗИЧНА ХАРАКТЕРИСТИКА ПОЛТАВСЬКОГО ПСГ

1.1 Географічне положення та історія створення об'єкта.

Полтавське підземне сховище газу (ПСГ) є важливим елементом газотранспортної системи України, оскільки забезпечує накопичення природного газу та його подачу в періоди підвищеного споживання. В Україні підземні сховища газу переважно створювались на базі виснажених газових родовищ, що є ефективним як з економічної, так і з геологічної точки зору. Такий підхід широко застосовується і в інших країнах світу (*Evans, Chadwick, 2009*). Завдяки цьому вдається використовувати вже сформовані природні резервуари для зберігання газу та забезпечувати стабільність роботи газотранспортної системи (*Iwaszczuk et al., 2022*).

Полтавське ПСГ розташоване в межах Петрівсько-Чутівської зони антиклінальних складок Дніпровсько-Донецької западини (див. Рис.1.1). Дана геоструктура характеризується значною потужністю осадових відкладів, складною будовою та високим ступенем нафтогазоносності, що робить її сприятливою для створення підземних сховищ газу (*Demchuk et al., 2025*). Саме тому більшість газових родовищ і підземних сховищ України приурочені до цієї западини (*Iwaszczuk et al., 2022*).



Рис.1.1 - Розташування Дніпровсько-Донецької западини в межах тектонічної структури України

(Джерело: Гілецький та ін., 2008)

В адміністративному відношенні Полтавське ПСГ розташоване в межах Полтавської області, на території Диканського та Зіньківського районів. Воно знаходиться за 30 км від обласного центру — міста Полтава. Безпосередньо на площі родовища розташовані села Пругли, Винтикці, Занки, Волошкове, а поблизу — Опішня, Духовщина та Березівка. Така локалізація є досить зручною, оскільки територія добре освоєна, що полегшує доступ до об'єкта (*Звіт з геолого промислової характеристики УТГ, 2025*).

З точки зору транспортної доступності, район розташування ПСГ має розвинену мережу автомобільних доріг, які забезпечують зв'язок між населеними пунктами та виробничими об'єктами. Дороги мають як асфальтоване, так і ґрунтове покриття. Найближча залізнична станція знаходиться у місті Полтава, що дозволяє здійснювати транспортування обладнання та матеріалів (*Звіт з геолого промислової характеристики УТГ, 2025*).

У геоморфологічному відношенні територія Полтавського ПСГ належить до рівнинної частини лівобережної України. Рельєф представлений рівниною, яка поступово знижується у південно-західному напрямку. Разом з тим поверхня значною мірою розчленована ерозійними формами, такими як балки та яри. Абсолютні висоти змінюються в межах від 83–85 м у річкових долинах до 180–197 м на правобережжі р. Ворскла (*Звіт з геолого промислової характеристики УТГ, 2025*).

Гідрографічна мережа району представлена басейном річки Ворскли, яка є лівою притокою Дніпра. Окрім головної водної артерії, на території поширені її ліві притоки, серед яких Мерла, Ківжиха та Свинківка. Вони безпосередньо впливають на формування водного режиму території, забезпечують природний дренаж і відіграють важливу роль у формуванні гідрогеологічних умов району ПСГ. Розвинена гідрографічна мережа також визначає рівень залягання ґрунтових вод і впливає на загальні інженерно-геологічні умови об'єкта (*Звіт з геолого промислової характеристики УТГ, 2025*).

Кліматичні умови району розташування Полтавського ПСГ характеризуються як помірно континентальні, що є типовим для лісостепової зони України. Для цієї місцевості властиве тепле літо та порівняно м'яка зима з середньорічною температурою повітря на рівні +7 °С. У літній період температурні показники зазвичай коливаються в межах +23...+25 °С, тоді як у зимові місяці вони знижуються до –6...–8 °С при середній річній кількості опадів близько 450–500 мм (*Звіт з геолого промислової характеристики УТГ, 2025*).

Такі метеорологічні умови вважаються цілком сприятливими для функціонування об'єкта, оскільки відсутність критичних температурних перепадів мінімізує навантаження на наземні інженерні споруди та технологічне обладнання. Крім того, вітровий режим території має виражену сезонність: у теплу пору року переважають потоки західного та північно-західного напрямків, а в холодну — північно-східного. Окрему увагу при проектуванні та експлуатації технологічних

мереж приділяють глибині промерзання ґрунтів, яка в даному районі становить 0,7–1,5 м, що обов'язково враховується для забезпечення надійності заглиблених конструкцій та трубопроводів (*Звіт з геолого промислової характеристики УТГ, 2025*).

В економічному аспекті район робіт має виражену аграрну спрямованість: основними видами діяльності тут є тваринництво, а також вирощування зернових і технічних культур. Попри це, провідну роль у господарстві відіграє саме нафтогазова галузь. Висока промислова освоєність регіону зумовила створення розгалуженої мережі спеціалізованих підприємств та відповідної інфраструктури, що стало ключовою передумовою для подальшого технологічного розвитку об'єкта (*Коболев та ін., 2018; Demchuk et al., 2025*).

Саме поєднання вигідного географічного положення та наявної промислової бази заклало фундамент для перетворення виснаженого родовища на стратегічно важливий вузол газотранспортної системи. Історія становлення Полтавського ПСГ розпочалася в період активної розбудови мережі підземного зберігання в Україні, коли виникла гостра потреба у створенні надійних резервних потужностей для забезпечення стабільного газопостачання великих промислових центрів (*Chernova, 2014*).

Історія освоєння Полтавського об'єкта розпочалася задовго до його перетворення на підземне сховище, ставши результатом масштабних геофізичних пошуків середини ХХ століття. Вперше структуру було виявлено у 1952 році в мезозойських відкладах за допомогою сейсмічних досліджень тресту «Укрнафтогазгеофізика», а згодом ці дані підтвердили структурно-картувальним бурінням по відкладах київської світи палеогену.

У 1953 році стартувало пошукове буріння, спрямоване на промислову оцінку мезозойського комплексу, що призвело до отримання наприкінці 1954 року перших промислових притоків газу з байоського ярусу середньої юри (свердловини № 7, 8, 9). Це дозволило у 1957 році офіційно прийняти родовище на Державний

баланс із підтвердженими запасами юрського покладу в обсязі 1871 млн м³ (Роздорожний, 2015; Звіт з геолого промислової характеристики УТГ, 2025).

Подальше вивчення площі пов'язане з освоєнням глибших горизонтів. У 1964 році шляхом сейсморозвідки було уточнено будову кам'яновугільного комплексу, а в 1966 році фахівці отримали промисловий газ із серпуховських відкладів раннього карбону (свердловини № 29, 5).

Проте багаторічна активна експлуатація призвела до того, що на початку 1980-х років юрський поклад виснажився. Саме тоді було ухвалено стратегічне рішення про використання його геологічного об'єму для створення підземного сховища газу. Офіційним роком утворення ПСГ став 1987-й, а вже у 1989 році об'єкт було виведено на проектні показники об'єму (Chernova, 2014; Роздорожний, 2015).

Протягом 1984–1991 років на площі тривали інтенсивні бурові роботи з метою формування ПСГ. У 1990–1991 роках для уточнення деталей будови та оцінки можливостей розширення сховища була проведена додаткова розвідка, яка охопила крайню східну частину підняття. Важливо зазначити, що на певних етапах замість буріння проектних свердловин у західній та центральній частинах площі для вивчення структурного плану та уточнення трасування тектонічних порушень було обрано метод сейсморозвідки, що дозволило оптимізувати процес дослідження об'єкта (Коболев та ін., 2018; Фондові матеріали Укртрансгаз, 2025).

За результатами сейсмічних досліджень 1960-х та 1980-х років було встановлено, що структура значно розширюється у західному та північному напрямках (див. Рис.1.2) . Це призвело до виявлення самостійних структур на перикліналях підняття: Західно-Солохівської, Кирилівської, Устимівської та Покровської, а також картування Бакейського соляного штоку.

Результати робіт Східно-Української геофізичної експедиції 1997–2001 рр. остаточно підтвердили двоповерхову будову структури, де нижній поверх складають відклади девону та нижнього карбону, а верхній — від

верхньовізейських відкладів до мезозою включно (Роздорожний, 2015; Demchuk et al., 2025).

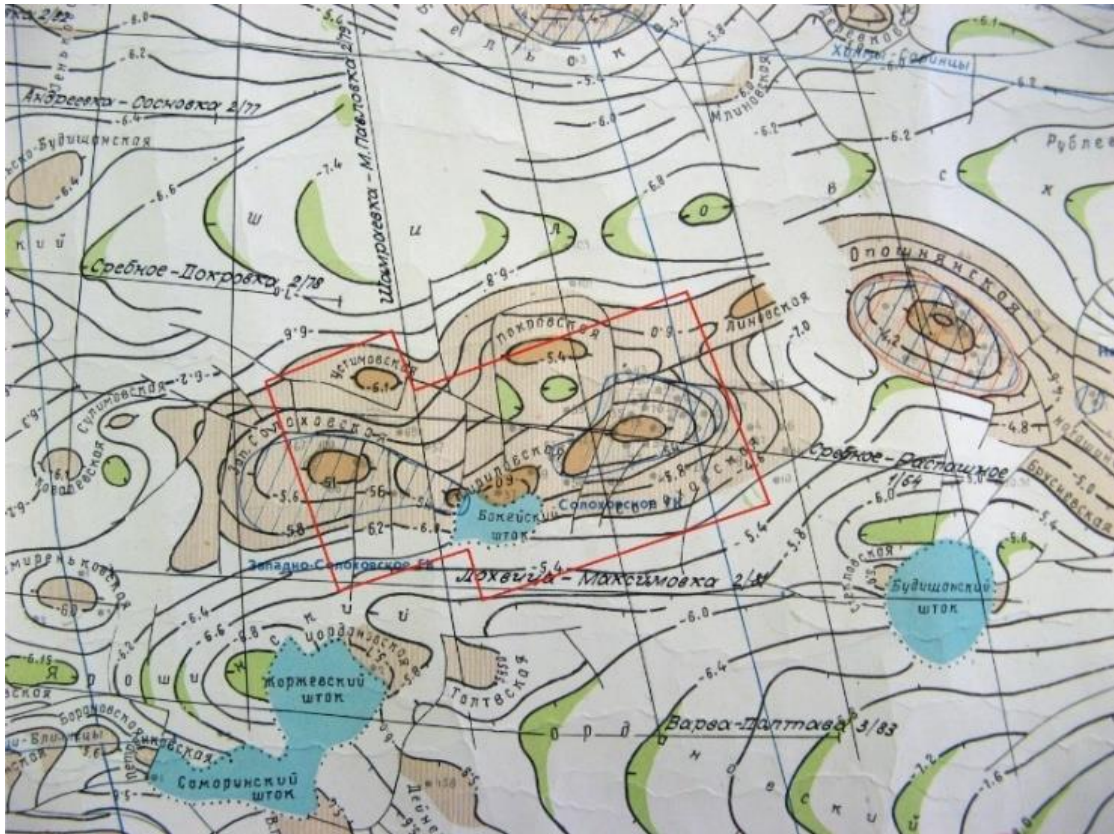


Рис. 1.2 - Схематична карта блокової будови Полтавського підняття та прилеглих територій

(Джерело: Звіт з геолого промислової характеристики УТГ, 2025)

Питанням структурно-геологічної побудови підняття в різні часи займалися такі дослідники, як Г. С. Брайлівський, В. Є. Орел, Б. П. Гоцький та Є. Є. Вороной. Усі автори сходилися на тому, що об'єкт є класичною куполоподібною брахіантикліналлю зі складною блоковою будовою, хоча їхні погляди на кількість блоків та трасування розривів дещо різнилися залежно від обсягу доступної на той час інформації.

Протягом 1993–2005 років відділом ПСГ інституту «УкрНДІгаз» було виконано фундаментальний аналіз усього масиву ГДС, буріння та сейсморозвідки, що дозволило деталізувати тектонічні особливості об'єкта (Коболев та ін., 2018).

Новий етап у дослідженні Полтавського ПСГ розпочався у 2006–2007 роках завдяки масштабній 3D-сейсморозвідці, яку виконала «Укрнафтогазгеофізика». Оскільки ці роботи охопили значну частину площі сховища, з'явилася можливість детальніше перевивчити відклади юрського комплексу. Фактично, саме ці нові сейсмічні дані дали змогу перейти до створення високоточних геологічних моделей і заклали базу для розробки «цифрових двійників», що зараз є головним інструментом для оптимізації роботи ПСГ. (*Demchuk et al., 2025; Звіт з геолого промислової характеристики УТГ, 2025*).

1.2 Геологічна будова та літолого-петрофізична характеристика району

Стратиграфія

Геологічний розріз Полтавського підняття відзначається значною стратиграфічною повнотою, охоплюючи відклади палеозойської, мезозойської та кайнозойської груп. Завдяки проведеним пошуковим, розвідувальним та експлуатаційним роботам вдалося детально вивчити товщу порід кам'яновугільної, тріасової, юрської, крейдової та палеоген-неоген-антропогенної систем. Цікавою особливістю є те, що в межах південно-західної перикліналі структурою зафіксовано штокове залягання девонської кам'яної солі, яку розкрила свердловина № 37.

Оскільки ПСГ було створене на базі виснаженого газового покладу в байоському ярусі середньої юри, у роботі зосереджено увагу на детальному описі саме верхньої частини розрізу — від тріасу до четвертинних відкладів включно. Це дозволяє зосередитися на характеристиках безпосередньо продуктивного об'єкта та відкладах, що залягають вище (*Фондові матеріали Укртрансгаз, 2025*).

ПЕРМСЬКА СИСТЕМА

На досліджуваній ділянці пермська система представлена виключно піщано-глинистими породами, які залягають на розмитій поверхні кам'яновугільних відкладів. Між цими системами чітко спостерігається різке кутове неузгодження. Сама товща має характерне строкате забарвлення, проте через відсутність будь-яких органічних залишків її вік довгий час вважався проблематичним. Сьогодні ми визначаємо його за непрямыми ознаками, спираючись на регіональні стратиграфічні зв'язки. Фондові дані підтверджують, що в межах Дніпровсько-Донецької западини та на околицях Донбасу ця товща лягає на різні горизонти пермі та карбону, а зверху її повсюдно перекривають базальні піщано-конгломератові шари тріасового віку. Саме через таке положення в розрізі прийнято вважати описуваний комплекс верхньопермським (*Фондові матеріали Укртрансгаз, 2025*).

Літологічно ця товща починається з базального шару, який за аналогією з Радченківським чи Сагайдацьким районами складається з піщано-глинистого матеріалу з включеннями дрібних уламків карбонатових вапняків та пісковиків. Цей горизонт фіксує межу розмиву та початок нового етапу накопичення осаdів. Вище по розрізу спостерігається часте перешаровування темно-коричневих і червоноколірних глин із зеленувато-сірими пісковиками. Примітно, що у верхній частині розрізу карбонатність порід зростає, аж до появи прошарків сильно піщанистих вапняків.

Загалом усю товщу можна розділити на дві частини. Нижня частина є переважно алеврито-глинистою; вона складена щільними різнокольоровими глинами з підпорядкованими прошарками алевролітів, які іноді переходять у дрібнозернисті пісковики. Потужність цього інтервалу становить близько 80–90 метрів. Верхня частина має більш виражений глинисто-піщанистий характер. Тут ми бачимо ритмічне чергування пластів потужністю від 2 до 6 метрів, де сірі пісковики змінюються цегляно-червоними піщаними глинами. Товщина верхнього

пакунку становить 45–55 метрів. Якщо говорити про загальну потужність верхньопермських відкладів, то в межах підняття вона зазвичай коливається від 138 до 150 метрів, хоча в окремих точках, як-от у районі Рунівщини, може досягати навіть 200 метрів (*Фондові матеріали Укртрансгаз, 2025*).

ТРИАСОВА СИСТЕМА

Тріасові відклади на цій ділянці лежать на верхній пермі, причому з помітними слідами розмиву та невеликим кутовим неузгодженням. За своїм складом це типовий континентальний тріас, який ми звикли бачити на північно-західних окраїнах Донбасу. Цікаво, що саме в Дніпровсько-Донецькій западині (і на нашому родовищі зокрема) потужність тріасу різко зростає. Це прямо натякає на те, що в той час територія ДДз прогиналася значно інтенсивніше, ніж сусідній Донбас чи Російська платформа.

З палеонтологією тут складно: знайти якісь рештки фауни майже неможливо, трапляються хіба що поодинокі остракоди та харові водорості. Тому розріз ділять на частини суто за «зовнішністю» порід, виділяючи три основні товщі: піщанисту, піщано-карбонатну та глинисту (*Фондові матеріали Укртрансгаз, 2025*).

Піщаниста товща — це самий низ тріасу. Вона залягає на розмитій пермі й добре простежується по всьому району ПСГ. В основному це зеленувато-сірі пісковики та піски з домішками вапна, слюди та піриту. У самому низу вони грубі, з галькою кварцу, яшми та кременю, але чим вище по розрізу, тим матеріал стає дрібнішим. Глини та алевроліти тут зустрічаються рідко, лише як тонкі прошарки. Потужність цієї пачки становить приблизно 110–126 м.

Далі йде **піщано-карбонатна товща**. Це такий собі «шаруватий пиріг» із частим чергуванням сірих алевролітів, пісковиків та строкатих (часто фіолетово-коричневих) глин. Головна фішка цієї частини — вапнякові жовна, які подекуди збираються в цілі прошарки. За всіма ознаками ці породи відповідають нижній

частині сребрянської свити середнього тріасу. Її потужність коливається від 54 до 74 м. (*Фондові матеріали Укртрансгаз, 2025*).

Замикає розріз **глиниста товща**. Як зрозуміло з назви, тут панують глини: цегляно-червоні, бурі або зеленуваті, буквально набиті дрібними карбонатними конкреціями. Між ними трапляються пласти сірих пісковиків та алевролітів товщиною від пари сантиметрів до трьох метрів. Ця частина розрізу теж корелюється із сребрянською свитою Донбасу, а її потужність на площі досить солідна — 209–236 м.

Окремо виділяють **протопівську свиту**. Вона складається з пачок світло-бурих глин та дрібних пісковиків. Якщо порівнювати її з попередніми відкладами, то тут розріз стає ще більш глинистим (*Фондові матеріали Укртрансгаз, 2025*).

ЮРСЬКА СИСТЕМА

Юрські відклади залягають на тріасових породах із помітною стратиграфічною перервою, через що нижній відділ юри в цьому розрізі повністю випав. Як і на більшій частині Дніпровсько-Донецької западини, тутешня юра представлена сіроколірними, а у верхній частині — строкатими піщано-глинистими породами. Ці відклади мають змішаний характер: тут перемішані як морські, так і континентальні утворення середнього та верхнього відділів системи.

Цікаво, що в межах нашої структури юрські пласти досить витримані — це стосується і їхнього складу (якщо не брати до уваги саму підосву), і товщини. Вони дуже чітко корелюються між собою за даними каротажу свердловин. Оскільки в розрізі є кілька надійних маркуючих горизонтів, ми використовуємо їх як основні репери, коли робимо структурні побудови чи карти. Це дуже зручно, бо такі пласти-орієнтири дають змогу точно «прив'язатися» до глибини та зрозуміти загальну будову ділянки (*Фондові матеріали Укртрансгаз, 2025*).

СЕРЕДНЯ ЮРА

Байоський ярус (J_2^{bs}).

На відкладах, які ми умовно відносимо до аалену, лежить серія порід, що чітко ділиться на дві частини: піщано-глинисту внизу та чисто глинисту зверху. За загальними ознаками тут виділяють два горизонти, які охоплюють увесь нижній байос і частину верхнього. Весь цей комплекс ми зазвичай описуємо як нерозчленований байос.

Якщо дивитися на склад, то це таке собі «брудне» чергування: сірі алевроліти, кварцові пісковики (від землястого до жовтуватого кольору) та темні, майже чорні глини, які зазвичай не шиплять від кислоти. Цікавий момент у підшві пачки: там у глинах трапляється кутааста галька кварцу, а пісковики перетворюються на справжні гравеліти з камінцями до пів сантиметра. До того ж і в глинах, і в пісковиках часто трапляються скупчення рослинних решток. Вище по розрізу, у середній частині, з'являються плитчасті глини з тонкими прошарками алевроліту, а на самому верху вже переважають звичайні сірі дрібнозернисті піски (*Фондові матеріали Укртрансгаз, 2025*).

З фауною тут все досить сумно — вона дуже бідна і невиразна. В основному це піщані форамініфери на кшталт *Protonina* чи *Ammodiscus*, якісь поодинокі ядра та мегаспори. Такий мізерний набір мікроорганізмів прямо натякає, що умови в морі тоді були специфічні: можливо, сильно стрибала солоність або не вистачало карбонату кальцію. Схожа картина спостерігається і в східній частині нашої западини, де зміна фацій завжди тягне за собою різку зміну складу форамініфер. Власне, через таку специфіку ці відклади й об'єднують у нерозчленований байос. Потужність усієї цієї пачки коливається в межах 52–76 метрів (*Фондові матеріали Укртрансгаз, 2025*).

Батський ярус (J_2^{bt}).

Проведення межі між байоським та батським ярусами ускладнюється літологічною однорідністю порід. Через це її неможливо виділити візуально за керном, тому стратиграфічне розчленування спирається виключно на фауністичні залишки. Водночас батський ярус поділяється на два горизонти (нижній і верхній), які вже мають між собою певні літологічні відмінності.

Нижній бат представлений переважно сірими, а подекуди навіть симпатичними блакитно-сірими глинами. Вони бувають алевритистими, мають виражену шаруватість і іноді трохи «шиплять» від кислоти, бо містять вапняк. Потужність цієї частини розрізу коливається в межах 42–72 метрів (*Фондові матеріали Укртрансгаз, 2025*).

З верхнім батом ситуація дещо цікавіша, бо там уже йде справжнє чергування порід. Це сірі або зеленуваті безкарбонатні глини з мікрошаруватістю (там буквально міліметрові прошарки світлого алевроліту), а також слюдисті алевроліти, піски та навіть доломітизовані вапняки. В алевролітах часто можна знайти скупчення обвуглених залишків рослин. До речі, з викопних рештків тут трапляються зуби риб, відбитки листя та поодинокі остракоди. Товщина верхнього бату трохи менша — десь від 45 до 58 метрів (*Фондові матеріали Укртрансгаз, 2025*).

Верхній відділ.

Що стосується межі між середнім і верхнім відділами юри в межах ДДз, то вона проводиться досить умовно — зазвичай за «маяк» беруть покрівлю тих самих мікрошаруватих глин верхнього бату. На нашій площі верхня юра представлена в повному обсязі й складається переважно з морських відкладів.

На відміну від бідних на органіку нижніх пластів, тут нам пощастило: породи буквально напхані різною мікро- та макрофауною. Якщо розбирати по ярусах, то розріз включає келовей, оксфорд та кімеридж. (*Фондові матеріали Укртрансгаз, 2025*).

Келовейський ярус (J_3^{kl})

Келовей у нас ділиться на два комплекси, які реально відрізняються і за тим, як вони виглядають, і за умовами, у яких вони утворювалися.

Нижній комплекс — це така собі «мішанка» із сірих глин різних відтінків (від світлих до майже чорних). Вони слюдисті, часто записочені й переслаюються з дрібними кварцовими пісковиками. Подекуди в цій товщі трапляються пласти міцного вапняку, бурі сидерити та вкраплення піриту. Що цікаво: у самій підшві повно обвуглених залишків рослин і навіть трапляються тоненькі прошарки бурого вугілля. Це явна ознака того, що спочатку умови тут були континентальні. Але чим вище по розрізу, тим більше стає карбонатів і з'являється справжня морська фауна: амоніти, белемніти та форамініфери. Саме по них ми й визначаємо, що це нижній келовей. Потужність цієї частини — десь 22–32 метри (*Фондові матеріали Укртрансгаз, 2025*).

Верхній комплекс (це середній та верхній келовей) виглядає вже інакше. Це переважно сірі або блакитно-сірі піщанисті глини, які дуже сильно вапнисті. Головна їхня фішка — величезна кількість макро- та мікрофауни. Тут амоніти та форамініфери зустрічаються на кожному кроці. Цей комплекс значно тонший за нижній — всього від 8 до 14 метрів.

Оксфордський ярус (J_3^{oxf})

Вище за келовей у нас іде оксфордський ярус. Межа між ними проводиться дуже чітко завдяки класному літологічному реперу — це пачка попелясто-сірих вапняків із прошарками блакитно-зеленої глини. Хоч цей шар і тонкий (усього 4–5 метрів), він ідеально «відбивається» на каротажних діаграмах і добре простежується по всьому району, тому ми його використовуємо як головний опорний горизонт для всіх структурних побудов.

Сам оксфорд ми ділимо на три частини, як це зазвичай і прийнято для цього ярусу:

1. **Нижній оксфорд:** Тут переважають зеленувато-блакитні карбонатні глини. Місцями вони трохи записочені, іноді трапляються рідкісні прошарки глинистого сидериту. З фауни тут стандартний набір — амоніти та форамініфери. Потужність невелика, десь від 5 до 10 метрів.
2. **Середній оксфорд:** Це вже солідна товща — від 80 до 100 метрів. Склад схожий (ті ж самі зеленувато-блакитні глини), але ближче до верху починають з'являтися часті тонкі прошарки пісковиків та детритусових вапняків.
3. **Верхній оксфорд:** Тут уже йде справжнє чергування порід. До глин додаються алевроліти, пісковики та тонкі вапняки, які буквально забиті битою черепашкою (детритом). Крім амонітів, тут уже можна знайти й устриць. Товщина цієї частини становить приблизно 43–55 метрів.

Якщо підсумувати, то загальна потужність оксфордського ярусу на нашій площі досить пристойна — від 128 до 165 метрів (*Фондові матеріали Укртрансгаз, 2025*).

Кімериджський ярус (J_3^{km}).

На оксфорд зверху лягає кімеридж, і межу між ними знайти взагалі не проблема. Там у підосві є характерний шар зеленого кварцово-глауконітового пісковика — він дрібнозернистий, пухкий і в полі відразу кидається в очі. Саме його ми й вважаємо кордоном між ярусами.

Нижня частина кімериджу — це чергування плямистих глин (блакитних та буро-коричневих) із вапнистими алевролітами й дрібними пісковиками. Подекуди трапляються зовсім тонкі прошарки вапняку, набитого всякою органікою. Завдяки форамініферам, які сидять у глинистих прошарках, ми точно знаємо, що це нижній під'ярус.

А далі розріз стає зовсім яскравим і строкатим. Починається товща глинистих алевролітів та піщанистих глин, де кольори просто на будь-який смак: сургучно-червоні, бурі, жовтуваті та блакитні. У всій цій масі трапляються пласти блакитно-сірих (іноді теж кольорових) пісковиків. Загалом вся ця штука має потужність від 83 до 151 метра (*Фондові матеріали Укртрансгаз, 2025*).

КРЕЙДОВА СИСТЕМА

Крейда в нашому районі є практично всюди. Вона лягає на розмиту поверхню верхньоярських відкладів. Тут у нас представлені обидва відділи — і нижній, і верхній. Якщо говорити про товщину, то на ПСГ цифри досить серйозні: загальна потужність коливається від 200 до 480 метрів.

Нижній відділ (Cr₁).

Нижня крейда представлена пачкою світло-сірих пісковиків різного розміру, які перемішуються із сірими, а іноді зовсім білими піщанистими глинами. Цікаво, що чим вище ми йдемо по розрізу, тим більше стає глин. Причому і вони, і пісковики у верхній частині помітно темнішають — це все через велику кількість вуглистої органіки. Власне, саме в цих темних вуглистих глинах палеонтологи знайшли спори, за якими й підтвердили, що це саме нижньокрейдові відклади. Товщина всієї цієї товщі на ділянці складає приблизно від 50 до 90 метрів (*Фондові матеріали Укртрансгаз, 2025*).

Верхній відділ (Cr₂).

Верхня крейда на нашій площі збереглася набагато краще, ніж на сусідніх ділянках (типу того ж Радченкового). Тут розріз представлений більш повно, і в ньому спокійно можна виділити сеноманський ярус, а також увесь сенон-туронський комплекс (*Фондові матеріали Укртрансгаз, 2025*).

Сеноманський ярус (Cr_2^{cm})

Сеноман у нас тут розпадається на дві пачки. Нижня — піщана: це пухкі зеленувато-сірі пісковики (кварц-глауконітові) з прошарками кварцитів та піщанистих глин у самому низу. Верхня пачка вже мергельна, це десь 10–15 метрів світло-сірого мергелю, у якому багато великих зерен глауконіту та дрібної гальки кременю. До речі, і в пісках, і в мергелях знайшли форамініфери, які чітко вказують на сеноманський вік. Загалом ця вся товща має потужність від 40 до 60 метрів (*Фондові матеріали Укртрансгаз, 2025*).

Сенон-турон (Cr_1^{sn+tur})

Сенон-туронський комплекс — це, по суті, класична біла крейда і світло-сірі мергелі, у яких часто трапляються включення сірого кременю. Що цікаво, повнота розрізу тут прямо залежить від рельєфу: на підняттях (у склепіннях) залишився сенон-турон і тільки шматочок сантону, а от на занурених ділянках ситуація значно краща — там і сантон на місці, і навіть кампанський ярус частково зберігся.

До речі, в керні реально знайшли фауну турону, кон'яського та інших ярусів, так що з віком тут усе підтверджено. Потужність усієї цієї товщі досить солідна — вона гуляє в межах 150–275 метрів (*Фондові матеріали Укртрансгаз, 2025*).

ПАЛЕОГЕН-НЕОГЕНОВА СИСТЕМА

Щодо палеогену, то він тут представлений трьома основними ярусами: канівсько-бучакським, київським та харківським. Це, по суті, класичний набір для нашого регіону.

Канівсько-бучакські відклади лягають на верхню крейду трансгресивно, тобто море свого часу знову насунулося на старі породи після перерви. За складом це переважно дрібні кварцо-глауконітові піски, у яких трапляються прошарки

піщанистих глин та окремі лінзи твердих кременистих пісковиків. У самому низу, прямо в підшві пісків, лежить шар галечнику з дрібних камінців кременю та кварцу — це така собі базальна верства, з якої все починається. Товщина цих відкладів на ділянці зазвичай коливається від 27 до 47 метрів (*Фондові матеріали Укртрансгаз, 2025*).

Київський ярус залягає на бучакських пісках повсюдно і, що важливо, без жодних перерв (згідно). Весь цей ярус ми ділимо на три частини: низ із фосфоритами, мергелиста середина та глинистий верх.

Нижній горизонт — це такі собі піщанисті мергелі блакитно-сірого кольору. Його головна фішка — пласт фосфоритових конкрецій у самому низу. Це взагалі залізобетонний літологічний репер, за яким київський ярус впізнають одразу.

Далі йде середній горизонт. Це дуже однорідні, добре відмулені блакитно-сірі мергелі. Там зазвичай повно решток мушлів, їх реально багато.

А от верхній горизонт зустрічається не всюди, він розвинений обмежено. Там уже йдуть сірувато-зелені, дуже щільні шаруваті глини. Вони іноді бувають піщанистими, але, на відміну від мергелів, зовсім не містять карбонатів (*Фондові матеріали Укртрансгаз, 2025*).

Загальна потужність київського ярусу невелика — всього від 18 до 27 метрів.

Харківський ярус складений товщею зеленувато-сірих пісків, переважно дрібно- та середньозернистих, які зазвичай мають глауконітовий склад. Усередині цієї товщі трапляються прошарки глин та пухких пісковиків на глинистому або карбонатному цементі. Якщо дивитися на розріз, то пісковики та глини частіше зустрічаються в його нижній частині, тоді як піски переважають ближче до верху. Іноді все це завершується пачкою дуже щільних темно-зелених глин товщиною від пів метра до шести метрів. Загалом харківські відклади досить потужні — їхня товщина коливається від 45 до 100 метрів.

Що стосується неогену, то ці піски та глини залягають на харківському ярусі й розповсюджені по всій площі, за винятком річкових долин — найбільше вони

розвинені саме на вододілах. У нижній частині неогенова товща представлена світло-сірим дрібнозернистим кварцовим піском, який має специфічні жовтуватобурі смуги. Вище по розрізу піски змінюються строкатими піщанистими глинами, у яких часто зустрічаються карбонатні та залізисті включення. Потужність неогенових відкладів невелика і в середньому становить близько 20–30 метрів (*Фондові матеріали Укртрансгаз, 2025*).

ЧЕТВЕРТИННА СИСТЕМА

Ну і наостанок — четвертинні відклади. Вони в нашому районі розповсюджені буквально повсюдно. На високих ділянках плато вони залягають прямо на неогенових строкатих глинах, а от у долинах річок чи по схилах балок вони «наповзають» на більш давні шари — неогенові піски або навіть на харківський ярус.

Якщо розбирати розріз на плато, то в самій основі четвертинної системи лежать червоно-бурі глини з карбонатними конкреціями, а вже над ними йде серія лесів та суглинків. Схили ж долин сформовані делювіальними та алювіальними відкладами, які, власне, і створюють річкові тераси та заплави.

Товщина цих утворень місцями досить солідна — на окремих ділянках вона сягає 45–55 метрів (*Фондові матеріали Укртрансгаз, 2025*).

ТЕКТОНІКА

У тектонічному плані наше підняття входить до Петрівсько-Чутівської зони антиклінальних складок, яка є частиною Диканьсько-Солохівського валу в самому центрі Дніпровсько-Донецької западини. Зі сходу до нього підходить Диканьська структура, але вони розділені невеликою сідловиною шириноюдесь півтора кілометра. Сама Диканька сидить приблизно на 50 метрів глибше (*Фондові матеріали Укртрансгаз, 2025*).

Якщо дивитися на структуру по покрівлі байоського пласта-колектора, то це така собі складноблокова асиметрична брахіантикліналь, витягнута на захід та північний захід (див. Додаток Б). Вона досить велика: за ізогіпсою мінус 720 метрів її розміри складають приблизно 12,8 на 4,5 кілометри, а висота підняття — близько 50 метрів.

Усе підняття розсікає великий скид (порушення 2-2), який ділить структуру на дві частини: солідну східну та значно меншу західну. Площина цього розлому падає на захід під кутом 50–55 градусів. Цікаво, що амплітуда скиду зростає в бік південного заходу: якщо на початку це всього 5–10 метрів, то далі розрив сягає вже 50 метрів. Східна частина, де якраз і облаштували підземне сховище газу, — це класична брахіантиклінальна складка з широким плоским склепінням. Вона асиметрична: південно-західне крило зовсім полого, з кутом нахилу всього градус-півтора, а от північно-східне вже крутіше — до 2,5 градусів. Перикліналь тут теж дуже довга і похила, розтягнулася майже на 5 кілометрів (*Фондові матеріали Укртрансгаз, 2025*).

Цю східну частину розривні порушення додатково розбили ще на чотири блоки: північний, центральний, південний та західний. Наприклад, північний блок займає північно-східне крило структури. Його межею на півдні є досить великий розлом 3-3, нахилений на північ. За даними свердловин, амплітуда зміщення байоського колектора по цьому розлому коливається в межах 30–55 метрів (Див. Додаток А)

На заході цей блок обмежується розломом 2-2. За розмірами він невеликий — десь 1,1 на 2,5 км, а його висота сягає приблизно 50 метрів. Далі йде Центральний блок, і тут ситуація цікавіша: він затиснутий між скидами 3-3, 4-4 та 2б-2б, утворюючи справжній грабен. Один із цих розломів (4-4) досить крутий, падає на північ під кутом 70–75°, а зміщення пласта по ньому становить 20–30 метрів. Через таку будову верхівка цього блоку виявилася на 10–12 метрів нижчою за сусідні ділянки.

Найбільшим є Південний блок, який займає майже дві третини всієї східної площі підняття. Він обмежений розломами 4-4 та 26-26, має солідні розміри (6,7 на 2,6 км) і висоту біля 50 метрів. Поруч розташований Західний блок, який за формою нагадує напівсклепіння. Його «обрізають» розломи 2-2 та 26-26. Амплітуда зміщення по 26-26 неоднакова: у центрі це 30 метрів, а на південному крилі — всього 10. Оскільки із заходу цей блок екранується скидом 2-2, тут утворилася ідеальна структурно-тектонічна пастка для газу (*Фондові матеріали Укртрансгаз, 2025*).

Дуже важливий момент для роботи нашого ПСГ: хоча східна частина і розбита розломами, вони не розривають гідродинамічний зв'язок між блоками. Це пояснюється просто — амплітуда зміщення порід менша за товщину самого байоського пласта-колектора, тому газ може спокійно мігрувати між ними.

Якщо порівнювати, то західна частина підняття значно менша за східну (десь 4,6 на 3,6 км). Вона виглядає значно «пом'ятою»: крила сильно деформовані, а закінчується вона структурним носом. Тут теж є пастки, але вони дрібні та ізольовані одна від одної. Ця частина також ускладнена грабенем (розломи 2-2 та 2а-2а), де амплітуда зміщення сягає 25–30 метрів. До речі, біля свердловини 81 є припіднятий блок, куди цілком може перетікати флюїд із північної частини.

Крім цих основних розломів, на площі повно дрібних «опіряючих» тріщин, але на роботу сховища вони практично не впливають. Майже всі ці скиди утворилися ще до палеогену — вони прорізають усі нижні шари й «затиhaють» лише біля підшви канівсько-бучацького ярусу (*Фондові матеріали Укртрансгаз, 2025*).

ФІЗИКО-ЛІТОЛОГІЧНА ХАРАКТЕРИСТИКА БАЙОСЬКОГО ПЛАСТА-КОЛЕКТОРА

На цій ділянці розвідка чітко підтвердила газовий поклад, який сидить у байоських пісковиках середньої юри (див. Додаток А). Колектор тут реально

класний і стабільно простежується по всій площі. Покрівля пласта залягає на глибинах від 851 до 956 метрів, а його робоча товщина зазвичай становить близько 60–70 метрів. Хоча є свої нюанси: у деяких місцях вона доходить до 80 метрів, а там, де проходять тектонічні скиди, пласт може «стискатися» до 20–30 метрів. На каротажі цей горизонт видно неозброєним оком завдяки дуже чіткій аномалії ПС, тож знайти його в розрізі не проблема (*Фондові матеріали Укртрансгаз, 2025*).

За складом це переважно кварцові піски — від дрібних до зовсім грубих. Вони дуже пухкі та слабо зцементовані, тому майже не тримаються купи. Якщо дивитися на розріз знизу вгору, то в самому низу пласта взагалі трапляється гравій та гравеліти, а середина розрізу перебивається тонкими шаруватими глинами. Ближче до верху пісок стає більш однорідним і дрібнозернистим.

Щодо фізики, то пористість тут переважно хороша — в середньому по покладу виходить близько 27%, а проникність сягає 3 Дарсі. Колектори переважно чисті, хоча іноді трапляються більш глинисті прошарки. Оскільки пісок дуже рихлий і не тримає форму, при видобутку він починає «летіти» разом із газом, тож на свердловинах обов’язково доводиться ставити протипіскові фільтри. А щоб газ нікуди не втік, зверху поклад надійно притиснутий товстою «кришкою» зі щільних монтморилонітових глин верхнього байосу та бату потужністю від 58 до 77 метрів (*Фондові матеріали Укртрансгаз, 2025*).

1.3 Аналіз попередніх геолого-геофізичних досліджень та стан вивченості теми

На сучасному етапі розвитку нафтогазової геофізики відбувається помітний технологічний перехід. Якщо раніше основним завданням вважалася побудова якісної статичної 3D-моделі, яка фіксувала стан покладу на конкретну дату, то сьогодні фокус уваги змістився до концепції цифрових двійників. Це не просто

сучасний тренд, а спроба створити динамічну систему, яка еволюціонує разом із родовищем та оновлюється в режимі реального часу (*Grieves, 2014*).

Проте, незважаючи на значний потенціал, виникає фундаментальна проблема верифікації: як довести, що ця складна математична конструкція адекватно відображає реальні пластові умови і їй можна довіряти при прийнятті управлінських рішень.

Концептуальні засади цього підходу були закладені ще Майклом Грівзом, який понад десять років тому описав ідею «віртуального прототипу», що супроводжує складний об'єкт протягом усього його життєвого циклу. І якщо в машинобудуванні такі рішення адаптувалися відносно швидко завдяки чітко заданим параметрам деталей, то в геофізиці процес впровадження виявився значно складнішим (*Grieves, 2014*).

Природні об'єкти — це глибокозалегаючі пласти зі складною анізотропією, макро- та мікронеоднорідністю і високим ступенем невизначеності. Тому повноцінне використання цифрових двійників у нашій галузі стало можливим лише нещодавно, із суттєвим зростанням обчислювальних потужностей.

Визначальну роль у реалізації цих підходів відіграла конвергенція інформаційних (IT) та операційних (OT) технологій. Як зазначає А. Шаджі Джордж, саме ця інтеграція дозволила перейти до створення «живих» моделей. Замість дискретного завантаження даних фахівці отримують можливість працювати з постійним потоком інформації від датчиків тиску, температури та дебітів.

Відповідно, модель трансформується зі статичної схеми на активний інструмент, що відображає реальні фізичні процеси: гідродинаміку флюїдів, перерозподіл пластової енергії тощо. Це забезпечує якісно інший рівень контролю за об'єктом (*George, 2024*).

Особливої уваги заслуговує питання інтеграції великих масивів даних. У дослідженні Ф. Тао та співавторів запропонована архітектура двійника, що охоплює п'ять вимірів: фізичний об'єкт, віртуальну модель, сервіси, дані та зв'язки

між ними. Автори акцентують увагу на проблемі «інформаційного розриву» — ситуації, коли накопичено терабайти даних, але їхня розрізненість не дозволяє побудувати єдину несуперечливу картину (*Tao et al., 2018*).

На практиці це призводить до розбіжностей між прогностичними показниками моделі та фактичною роботою свердловин. Ф. Тао доводить, що подолати цей розрив можна лише через глибоку аналітику та постійний зворотний зв'язок (*Tao et al., 2018*). Для підземних сховищ газу з їхніми жорсткими циклами закачування та відбору цей аспект є критично важливим.

Проте жодні обчислювальні алгоритми не здатні компенсувати недоліки петрофізичної основи. Дослідження С. А. Вижви, І. М. Безродної та С. І. Козіонові демонструють складність кореляційних зв'язків між геофізичними даними та результатами лабораторних вимірювань. Зокрема, у карбонатних породах із хаотичним розподілом порового простору використання спрощених петрофізичних моделей призводить до суттєвих похибок (*Вижва та ін., 2012*).

Отже, верифікація цифрового двійника вимагає не лише математичного узгодження параметрів, а насамперед глибокого розуміння мікроструктурних процесів у породі. Без цієї основи будь-яка комп'ютерна модель залишатиметься суто візуальною абстракцією.

Розвиток сучасних систем моделювання є логічним продовженням тривалої еволюції від 2D-картування до складних геостатистичних 3D-реалізацій, що детально описано в працях К. Є. Закревського та В. Л. Попова. На сьогодні акцент змістився від самого процесу побудови сітки до оцінки достовірності результатів. Ключовим викликом стає здатність моделі коректно інтегрувати сейсмічні дані та історію розробки, враховуючи властиві геологічному середовищу ризики та невизначеності (*Закревский та ін, 2021*).

Показовим для дослідження є досвід Abdo E., щодо впровадження цифрових двійників для моніторингу витоків газу на ПСГ. Автори підкреслюють, що через циклічну експлуатацію та нестабільні термобаричні умови ПСГ мають

значно складнішу динаміку, ніж традиційні родовища, що обмежує ефективність стандартних підходів. Для забезпечення адаптивності моделі італійські вчені застосували складні генетичні алгоритми оптимізації (*Abdo. et al., 2023*).

Підсумовуючи, створення достовірних цифрових двійників ПСГ гальмується складністю обробки багатовимірних геологічних ознак: традиційні алгоритми надто чутливі до якості вибірок і специфіки циклічної експлуатації сховищ. Тому найбільш перспективним рішенням є відмова від жорстких статистичних обмежень на користь методів непараметричної кластеризації.

Висновки до розділу: Підбиваючи підсумок першого розділу, можна стверджувати, що Полтавське ПСГ є стратегічно важливим об'єктом із високим ступенем геологічної вивченості, де унікальні фільтраційно-ємнісні властивості байоського колектора та надійна ізолююча покрівля створюють необхідні фізичні умови для впровадження динамічних моделей. Аналіз світового досвіду підтвердив актуальність переходу від статичних 3D-побудов до «цифрових близнюків» на основі ІТ/ОТ конвергенції, проте виявлений дефіцит методик для перевірки їхньої адекватності в реальному часі підкреслює необхідність пошуку нових підходів до верифікації моделей.

Таким чином, цей розділ стає фундаментальною основою для подальшого дослідження, надаючи необхідний геологічний контекст та методологічні засади для реалізації головної мети роботи — перевірки гіпотези подібності геофізичних характеристик об'єкта.

РОЗДІЛ 2. ТЕОРЕТИЧНІ ОСНОВИ СТВОРЕННЯ ЦИФРОВИХ БЛИЗНЮКІВ У ГЕОФІЗИЧНИХ ДОСЛІДЖЕННЯХ

2.1 Поняття цифрового близнюка та його роль у нафтогазовій галузі

Сучасний етап розвитку нафтогазової галузі характеризується суттєвою трансформацією підходів до моделювання підземних об'єктів — від традиційних статичних моделей до впровадження динамічних цифрових систем. Якщо класичні геологічні моделі дозволяли відображати стан родовища або підземного сховища газу лише на певний момент часу, то в умовах інтенсивної експлуатації та постійного надходження нових даних такий підхід уже не забезпечує необхідного рівня точності та адаптивності.

Актуальною альтернативою є концепція цифрового близнюка, яка, згідно з підходом, запропонованим Майклом Грівзом, передбачає взаємодію фізичного об'єкта та його віртуального представлення в межах єдиної інформаційної системи. У такій системі цифрова модель не є статичною, а постійно оновлюється на основі нових даних, що дозволяє відображати зміну стану об'єкта в режимі, наближеному до реального часу (*Grieves, 2014*).

Генезис поняття «цифровий близнюк» пов'язаний із розвитком інженерних дисциплін, зокрема методології управління життєвим циклом продукції. У початковому вигляді ця концепція передбачала створення віртуального аналога технічного об'єкта, який функціонує паралельно з фізичною системою на всіх етапах її існування — від проєктування до завершення експлуатації. Такий підхід дозволяв здійснювати моніторинг, аналіз і оптимізацію функціонування об'єкта на основі цифрових даних.

Подальший розвиток ця концепція отримала в дослідженнях NASA, де цифровий близнюк розглядається як інтегрована модель, що поєднує фізичні процеси, експериментальні дані та методи прогнозування технічного стану об'єкта.

У цьому підході особливий акцент робиться на здатності моделі прогнозувати поведінку системи та оцінювати її стан на основі аналізу даних (*Glaessgen & Stargel, 2012*).

Таким чином, концепція цифрового близнюка поступово еволюціонувала від інженерного інструменту до універсального підходу, який використовується для аналізу складних систем у різних галузях, зокрема й у нафтогазовій промисловості. У спрощеному вигляді цифровий близнюк можна розглядати як динамічну цифрову модель, що безперервно оновлюється на основі нових даних і відображає актуальний стан об'єкта.

У контексті дослідження геологічних об'єктів цей підхід набуває особливого значення, оскільки прямий доступ до підземного середовища є обмеженим або повністю відсутнім. Формування уявлень про стан пласта здійснюється опосередковано — на основі інтерпретації результатів геофізичних досліджень свердловин, даних буріння та показників промислової експлуатації. У таких умовах цифровий близнюк виступає як інструмент, що дозволяє об'єднати ці дані в єдину узгоджену систему.

Технологія цифрового близнюка забезпечує інтеграцію різнорідних джерел інформації, включаючи геологічні, геофізичні та експлуатаційні дані, і дозволяє оновлювати модель у процесі надходження нової інформації. Це створює передумови для більш точного аналізу стану об'єкта та прийняття обґрунтованих рішень щодо його експлуатації (*Tao et al., 2019*).

За своєю структурою цифровий близнюк включає три основні компоненти: фізичний об'єкт, його цифрову модель та канал обміну даними між ними. У нафтогазовій галузі фізичним об'єктом може виступати як окрема свердловина, так і більш складні системи — родовище або підземне сховище газу. Цифрова складова представлена моделлю, що описує геометричну будову, фільтраційно-ємнісні властивості та поточний стан об'єкта.

Канал передачі даних забезпечує зв'язок між фізичною системою та її цифровим аналогом, дозволяючи інтегрувати результати нових геофізичних досліджень, показники роботи свердловин та інші експлуатаційні дані (Jones et al., 2020). Взаємодія цих компонентів формує єдину систему (Див. Рис.2.1), у якій цифрова модель постійно уточнюється та використовується для аналізу і прогнозування стану об'єкта (Jones et al., 2020).

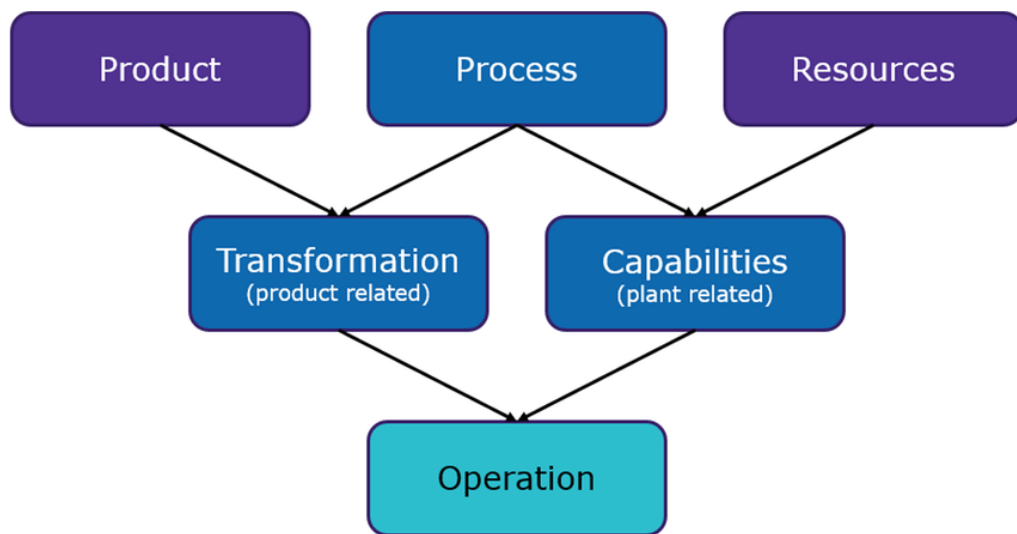


Рис.2.1 - Схема концепції отримання цифрового близнюка

(Джерело: Bamberg et al., 2021)

Позначення: Основними складовими моделі є: **Product** — продукт або об'єкт моделювання; **Process** — технологічний процес; **Resources** — ресурси системи; **Transformation** — процес трансформації або зміни стану об'єкта; **Capabilities** — функціональні можливості виробничої системи; **Operation** — кінцева операційна діяльність системи.

На рисунку 2.1 показано загальну концепцію функціонування цифрового близнюка виробничої системи. Як видно зі схеми, основою моделі є взаємозв'язок між продуктом, процесами та ресурсами, які разом формують єдину систему управління та аналізу даних. При цьому окремо виділяються процеси

трансформації та функціональні можливості системи, що в подальшому впливають на її роботу та ефективність.

Дана схема добре демонструє, що цифровий близнюк не є просто візуальною 3D-моделлю об'єкта. Насамперед це система, яка об'єднує різні типи даних, процеси та характеристики об'єкта в єдине цифрове середовище. Для підземних сховищ газу це особливо важливо, оскільки під час експлуатації постійно змінюються пластові тиски, режими закачування і відбору газу, а також гідродинамічний стан колектора. Саме тому ефективність цифрового близнюка безпосередньо залежить від якості та повноти даних, на основі яких формується модель об'єкта.

У геофізичному аспекті впровадження технології цифрового близнюка є особливо важливим, оскільки саме геофізичні методи виступають основним джерелом інформації про будову та властивості підземних резервуарів. Зокрема, сейсморозвідувальні дослідження дозволяють відтворити структурну будову пастки, визначити положення продуктивних пластів і виявити тектонічні порушення. У свою чергу, геофізичні дослідження свердловин дають змогу отримати більш детальну інформацію про літологічний склад порід, пористість, проникність і характер насичення флюїдами.

Разом з тим, суттєвою проблемою є те, що такі дані мають різний масштаб і часто є фрагментарними. У цих умовах цифровий близнюк можна розглядати як інструмент, який дозволяє об'єднати різноманітну геолого-геофізичну інформацію в єдину систему та використовувати її для подальшого аналізу і моделювання. Саме інтеграція даних є ключовою перевагою цього підходу, оскільки вона дає можливість отримати більш цілісне уявлення про об'єкт дослідження (*Tao et al., 2019*).

Ефективне функціонування підземних резервуарів і підземних сховищ газу (ПСГ) залежить не лише від правильної побудови їх геометричної моделі, а й від точності визначення петрофізичних параметрів. До основних із них належать

пористість, проникність, ефективна потужність пласта-колектора, а також літологічні характеристики покритишки і тип пастки. Саме ці параметри визначають фільтраційні процеси в пласті та впливають на поведінку газу в умовах експлуатації.

У межах цифрового близнюка ці характеристики не розглядаються як незмінні. Навпаки, вони постійно уточнюються на основі нових даних — як геофізичних, так і експлуатаційних. Це дозволяє моделі поступово адаптуватися до реального стану об'єкта та підвищувати точність прогнозів. Подібний підхід відповідає сучасним уявленням про цифрові моделі як динамічні системи, що розвиваються разом із накопиченням інформації (*Jones et al., 2020*).

У цьому контексті цифровий близнюк можна розглядати як логічне продовження класичного геолого-гідродинамічного моделювання. Якщо раніше модель створювалася як відносно статичний інструмент і використовувалася з мінімальними змінами, то сучасні підходи передбачають її постійне оновлення та уточнення (Див. Рис. 2.2).

Будь-які нові дані — результати геофізичних досліджень свердловин або зміни показників експлуатації — одразу враховуються в моделі. Завдяки цьому цифровий близнюк не втрачає актуальності та залишається максимально наближеним до фактичного стану підземного резервуара, що є ключовою умовою його ефективного використання (*Jones et al., 2020*).

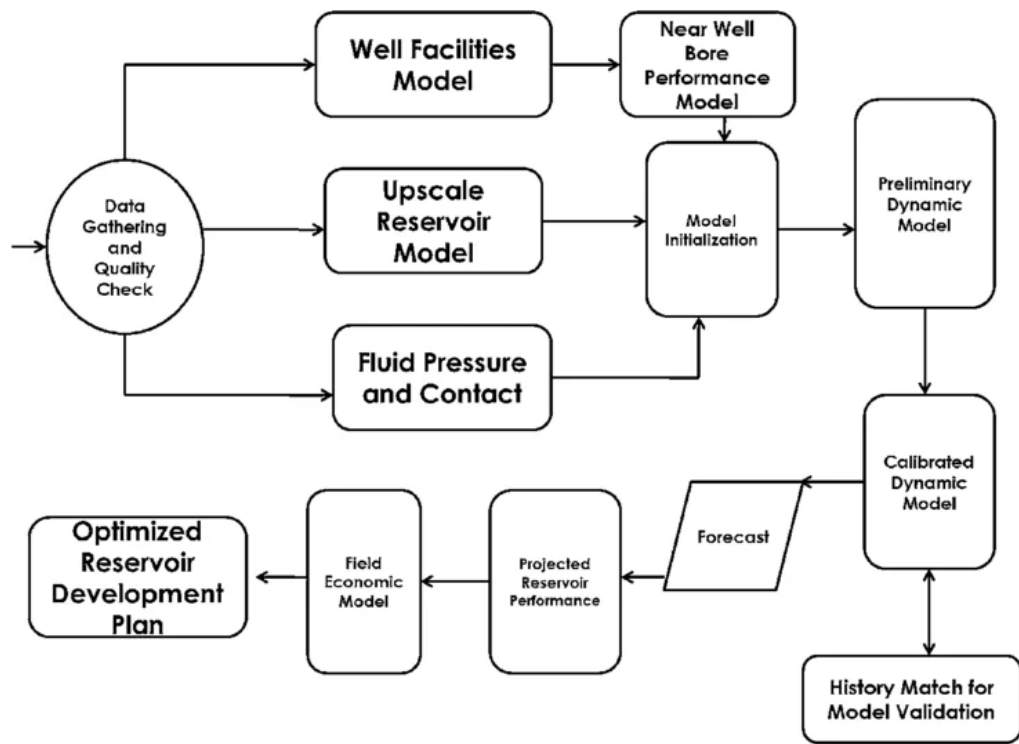


Рис.2.2 - Процес моделювання динамічного пласта

(Джерело: Sadri et al., 2019)

На рисунку 2.2 представлено загальну схему побудови та верифікації цифрової динамічної моделі родовища. Основними елементами схеми є збір даних та перевірка їх якості (**Data Gathering and Quality Check**), модель свердловинної інфраструктури (**Well Facilities Model**), модель роботи привибійної зони свердловини (**Near Well Bore Performance Model**), укрупнена модель резервуара (**Upscale Reservoir Model**), пластові тиски та флюїдні контакти (**Fluid Pressure and Contact**), ініціалізація моделі (**Model Initialization**), попередня динамічна модель (**Preliminary Dynamic Model**), відкалібрована динамічна модель (**Calibrated Dynamic Model**), адаптація та перевірка моделі за історією розробки (**History Match for Model Validation**), прогнозування (**Forecast**), прогнозна поведінка резервуара (**Projected Reservoir Performance**), економічна модель родовища (**Field Economic Model**) та оптимізований план розробки резервуара (**Optimized Reservoir Development Plan**).

Як видно зі схеми, процес створення цифрового близнюка починається зі збору, перевірки та систематизації вихідних геолого-геофізичних і промислових даних. Після цього формується модель резервуара, враховуються параметри привибійної зони свердловин, характеристики флюїдів і пластових контактів, а також створюється початкова динамічна модель об'єкта. Далі модель проходить етап калібрування та верифікації через зіставлення розрахункових параметрів із фактичними даними експлуатації. Такий підхід дозволяє поступово підвищувати точність моделі та адаптувати її до реальних умов роботи родовища.

Дана схема добре демонструє, що цифровий близнюк є не окремою моделлю, а безперервним процесом оновлення й аналізу даних. Для підземних сховищ газу це особливо важливо, оскільки режими закачування та відбору газу постійно змінюють стан колектора, пластові тиски та розподіл флюїдів у резервуарі. Саме тому ефективність цифрового близнюка безпосередньо залежить від якості геолого-геофізичних даних та точності їх інтеграції в єдину модель об'єкта.

Особливо актуальним використання цифрових близнюків є для підземних сховищ газу. Це пов'язано з тим, що ПСГ працюють у циклічному режимі, який супроводжується змінами пластового тиску, перерозподілом флюїдів у поровому просторі та можливим зміщенням газоводяного контакту. Такі процеси є типовими для експлуатації підземних резервуарів і потребують постійного контролю, оскільки вони безпосередньо впливають на ефективність та безпеку роботи сховища (*Mirfendereski et al., 2008*).

У цьому контексті застосування технології цифрового близнюка дозволяє більш системно підходити до аналізу таких процесів і вирішення практичних задач експлуатації. Зокрема, використання цифрового близнюка забезпечує:

- безперервний моніторинг змін пластових параметрів;
- прогнозування поведінки резервуара в різних режимах роботи;
- своєчасне виявлення та зниження технологічних ризиків;
- обґрунтування оптимальних режимів закачування та відбору газу.

Фактично цифровий близнюк виступає інструментом підтримки прийняття рішень, який дозволяє поєднати результати моделювання з фактичними даними експлуатації та підвищити ефективність управління роботою підземного сховища.

У межах теоретичного аналізу важливо також розмежувати поняття цифрового близнюка, цифрової моделі та цифрової тіні (Див. Рис. 2.3). Звичайна цифрова модель, як правило, є статичним відображенням об'єкта і не передбачає автоматичного оновлення даних. Цифрова тінь характеризується наявністю зв'язку з фізичним об'єктом, однак цей зв'язок є одностороннім — інформація передається лише від реального об'єкта до його віртуального представлення.

На відміну від цього, ключовою ознакою цифрового близнюка є двосторонній інформаційний обмін. У цьому випадку дані з реального об'єкта використовуються для постійного оновлення цифрової моделі, а результати моделювання можуть застосовуватись для коригування режимів роботи системи. Такий підхід дозволяє забезпечити більш точне відображення стану об'єкта та підвищити ефективність його експлуатації (Tao et al., 2019).

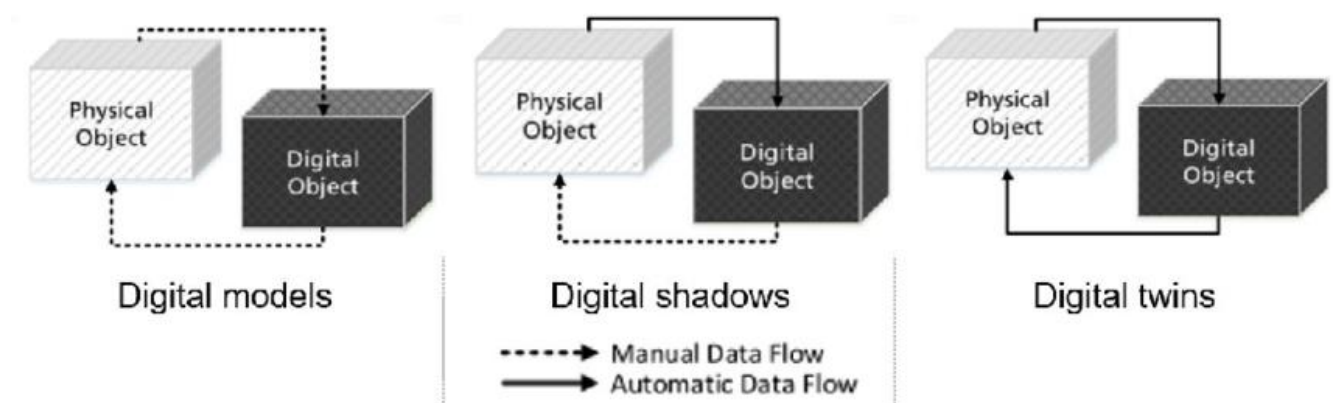


Рис.2.3 – Класифікація цифрових близнюків

(Джерело: Kritzinger et al., 2018)

На рисунку 2.3 показано різницю між цифровими моделями (**Digital models**), цифровими тінями (**Digital shadows**) та цифровими близнюками (**Digital twins**). Пунктирними стрілками позначено ручний обмін даними (**Manual Data Flow**), а

суцільними — автоматичний обмін даними (**Automatic Data Flow**) між фізичним (**Physical Object**) та цифровим об'єктом (**Digital Object**).

У випадку цифрових моделей обмін даними між фізичним і цифровим об'єктом практично відсутній або виконується вручну. Цифрові тіні вже передбачають автоматичну передачу даних від фізичного об'єкта до цифрового, однак зворотний зв'язок ще залишається обмеженим.

Найбільш повною системою є цифровий близнюк, де між фізичним і цифровим об'єктом відбувається постійний двосторонній автоматичний обмін даними.

У площині геофізичних досліджень ця концептуальна відмінність має принципове значення. Наприклад, побудова структурної моделі за даними сейсморозвідки сама по собі ще не дозволяє вважати її цифровим близнюком. Навіть у випадку інтеграції детальних результатів каротажу в тривимірну модель мова йде лише про цифрову копію об'єкта, якщо такі дані залишаються незмінними в часі.

Повноцінного статусу цифрового близнюка модель набуває лише тоді, коли геофізичні параметри постійно оновлюються, а сама модель використовується не тільки для опису, але й для аналізу та прогнозування поведінки об'єкта. Саме динамічність і можливість адаптації до нових даних є ключовими ознаками цього підходу (*Jones et al., 2020*).

У цьому контексті концепція цифрового близнюка може розглядатися як основа для перевірки гіпотези подібності геофізичних характеристик підземних об'єктів. Це означає, що увага зміщується від опису окремого пласта до порівняльного аналізу різних об'єктів з метою встановлення їх подібності за набором петрофізичних і геологічних параметрів.

Виявлення таких закономірностей створює передумови для застосування методів автоматизованого аналізу та класифікації, які дозволяють групувати

об'єкти за спільними характеристиками та використовувати отримані результати для подальшого моделювання (*Emerick & Reynolds, 2013*).

Підсумовуючи, можна зазначити, що в сучасній нафтогазовій геофізиці цифровий близнюк уже не є лише теоретичною концепцією, а виступає як практичний інструмент. Його основна цінність полягає у здатності об'єднувати різномірні дані, постійно уточнювати модель у процесі експлуатації та використовувати її для аналізу і прогнозування стану об'єкта. Для підземних сховищ газу такий підхід є особливо важливим, оскільки їх функціонування супроводжується складними і змінними внутрішньопластовими процесами (*Meza et al., 2024*).

2.2. Основні етапи та програмні засоби побудови цифрових моделей

У межах геофізичних досліджень створення цифрового близнюка є комплексним завданням, що поєднує обробку первинних даних, побудову математичних моделей та їх подальшу інтерпретацію (Див. Рис.2.5). У сучасних наукових підходах цифровий близнюк розглядається як віртуальне представлення реального геологічного об'єкта, яке на основі фактичних даних відтворює його будову, властивості та поведінку. Важливою особливістю такого підходу є його динамічність — модель не є фіксованою, а постійно уточнюється при надходженні нових даних (*Li et al., 2020*).

У випадку підземних сховищ газу концепція цифрового близнюка виходить за межі звичайної тривимірної моделі і набуває вигляду багатовимірної інформаційної системи. У її основі лежить аналіз даних, який дозволяє виявляти закономірності у великих масивах геолого-геофізичної інформації. Це відкриває можливість використовувати цифровий близнюк не лише для опису об'єкта, а й для класифікації пластів за подібністю їхніх характеристик (*Reichstein et al., 2019*).

Побудова цифрового близнюка починається з формування вхідного масиву даних. Для об'єктів ПСГ він включає стратиграфічні, літологічні та структурні параметри, такі як глибина залягання, потужність пластів, фазовий стан флюїдів і тип екранування. У задачах машинного навчання якість і повнота цих даних відіграють вирішальну роль, оскільки саме вони визначають точність подальшого моделювання. Тому на цьому етапі важливо не лише зібрати дані, а й забезпечити їх узгодженість і репрезентативність (*Liu et al., 2022*).

Наступним етапом є попередня обробка даних, яка є обов'язковою умовою побудови коректної моделі. Вона включає виявлення та усунення аномальних значень, роботу з пропущеними даними, а також нормалізацію параметрів для приведення їх до єдиного масштабу. Оскільки геологічні об'єкти описуються як кількісними, так і якісними ознаками, важливим завданням є їх приведення до формалізованого вигляду. Як показують дослідження, саме якість підготовки даних значною мірою визначає точність класифікації та надійність прогнозів (*Hastie et al., 2009*).

На основі підготовлених даних формується багатовимірний простір ознак, який є базою цифрового близнюка. У цьому просторі кожен пласт розглядається як окремий об'єкт із набором характеристик. Такий підхід дозволяє враховувати складні взаємозв'язки між параметрами. Наприклад, пласти можуть мати однакову глибину, але відрізнятися літологією або фазовим станом флюїду. Багатовимірний аналіз дозволяє врахувати ці відмінності одночасно, що підвищує точність класифікації (*Yang et al., 2022*).

Використання багатовимірного підходу є принциповою відмінністю сучасних методів аналізу від традиційних геологічних описів, де параметри часто розглядаються окремо. Перехід до аналізу взаємозв'язків між ознаками дозволяє більш повно описати складні геологічні системи та підвищити достовірність моделей (*Reichstein et al., 2019*).

Для класифікації пластів у даній роботі використовується метод дерев рішень як один із найбільш ефективних непараметричних підходів (Див. Рис. 2.4). Його перевага полягає у можливості побудови зрозумілої ієрархічної структури, де розподіл об'єктів відбувається за значеннями ознак. Такий підхід добре працює з неоднорідними даними та дозволяє враховувати складні поєднання геологічних параметрів. Як показано в дослідженнях класифікації резервуарів, методи на основі дерев рішень забезпечують високу точність виділення груп пластів із подібними характеристиками (Rahimi & Riahi, 2022).

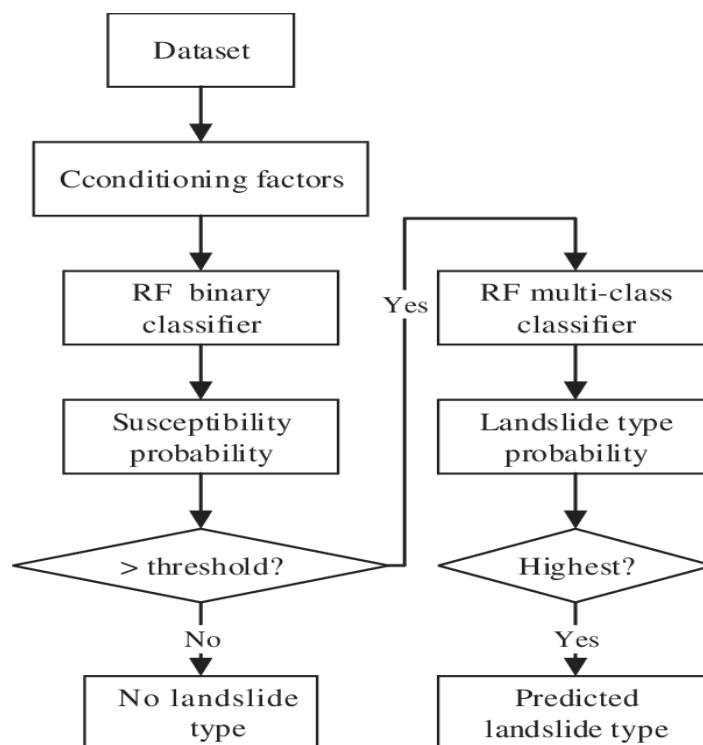


Рис.2.4 - Приклад структури дерева рішень
(Джерело: Huang et al., 2018)

На рисунку 2.4 представлено загальну схему роботи алгоритму класифікації даних. Основними елементами є набір даних (**Dataset**), фактори впливу (**Conditioning factors**), бінарний класифікатор (**RF binary classifier**), багатокласовий класифікатор (**RF multi-class classifier**), ймовірність належності до певного класу

(Susceptibility probability, Landslide type probability) та кінцевий прогнозований тип об'єкта **(Predicted landslide type)**.

Як видно зі схеми, на початковому етапі аналізуються вихідні дані та фактори, які впливають на формування певного процесу. Далі виконується бінарна класифікація, де система визначає, чи перевищує ймовірність заданий поріг. Якщо умова не виконується, об'єкт не належить до досліджуваної групи. Якщо ж поріг перевищено, запускається багатокласова класифікація, у межах якої визначається найбільш ймовірний тип об'єкта.

На кожному кроці обирається така ознака, яка дозволяє найкраще розділити об'єкти на більш однорідні групи. У результаті формується дерево, де кожен вузол відповідає певній умові, а кінцеві гілки — окремим класам об'єктів.

Отримана структура дозволяє не лише виконувати класифікацію, а й зрозуміти, які саме параметри найбільше впливають на розподіл пластів. Це особливо важливо для геологічних задач, де необхідно не просто отримати результат, а й пояснити його з фізичної точки зору (*Rahimi & Riahi, 2022*).

Після побудови моделі виконується її навчання та перевірка якості. Для цього набір даних поділяється на навчальну та тестову вибірки. Такий підхід дозволяє оцінити, наскільки модель здатна працювати з новими даними, які не використовувалися під час навчання. Це є важливим етапом, оскільки дає змогу уникнути перенавчання і забезпечити надійність результатів (*Kuhn & Johnson, 2013*).

Наступним етапом є інтерпретація отриманих результатів. У цьому випадку важливо не лише визначити класи, а й зрозуміти, як вони пов'язані з геологічними характеристиками пластів. Тобто результати моделі повинні бути узгоджені з реальними геологічними процесами.

У сучасних дослідженнях підкреслюється, що інтерпретованість моделей є ключовою умовою їх застосування в науках про Землю. Це дозволяє поєднати результати машинного навчання з фізичним змістом процесів, що відбуваються у підземних резервуарах (*Reichstein et al., 2019*).

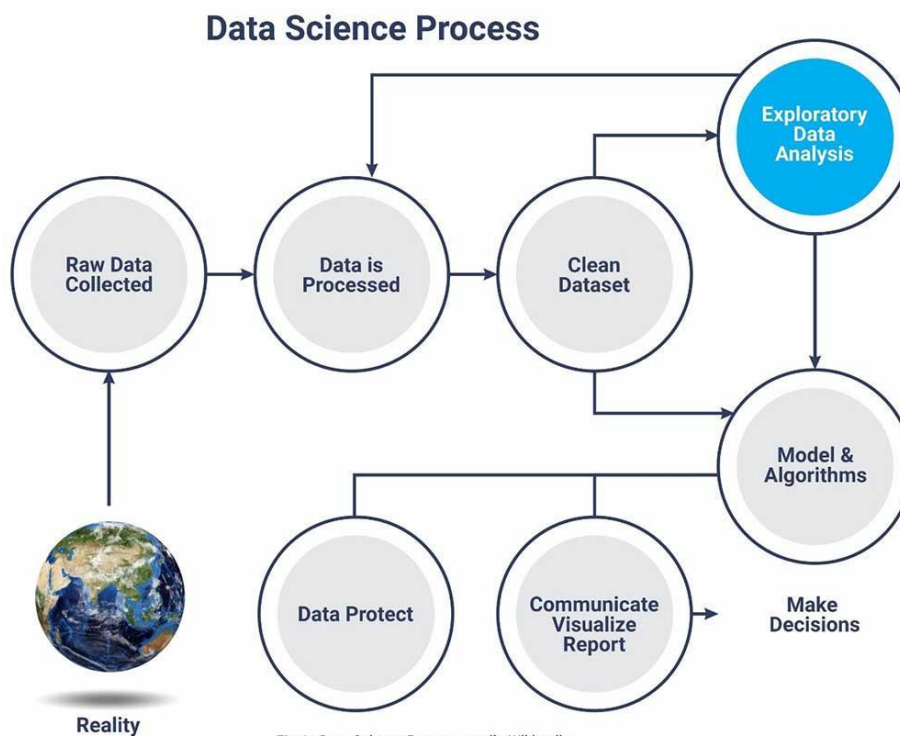


Рис.2.5 - Схема процесу побудови цифрової моделі на основі аналізу даних
(Джерело: Corrales et al., 2015)

На рисунку 2.5 представлено загальну схему процесу аналізу даних (**Data Science Process**). Основними етапами є збір початкових даних (**Raw Data Collected**), обробка даних (**Data is Processed**), формування очищеного набору даних (**Clean Dataset**), розвідувальний аналіз даних (**Exploratory Data Analysis**), побудова моделей та алгоритмів (**Model & Algorithms**), захист даних (**Data Protect**), візуалізація та представлення результатів (**Communicate Visualize Report**) і прийняття рішень (**Make Decisions**). Початковим джерелом інформації у схемі виступає реальний об'єкт або середовище (**Reality**).

Як видно зі схеми, процес аналізу даних складається з кількох послідовних етапів — від збору та первинної обробки інформації до побудови моделей і прийняття рішень. Спочатку формується підготовлений набір даних, після чого виконується їх аналіз для виявлення закономірностей і взаємозв'язків між параметрами.

Уже на основі оброблених даних створюються моделі та алгоритми для подальшої інтерпретації результатів. На мою думку, дана схема добре показує, що ефективність цифрової моделі напряду залежить від якості підготовки даних. У геофізичних дослідженнях це особливо важливо через велику кількість різномірної геолого-геофізичної інформації, тому саме етапи обробки та систематизації даних є основою для подальшого використання кластер-аналізу та мереж Кохонена.

Особливу увагу в роботі приділено перевірці гіпотези про подібність геофізичних характеристик досліджуваних пластів. Використання методів машинного навчання дозволяє перевести порівняльний аналіз об'єктів у формалізований вигляд, що дає змогу більш об'єктивно виділяти групи пластів із подібними параметрами. Такий підхід дозволяє вийти за межі простого опису окремих об'єктів і зосередитись на виявленні стійких закономірностей між їхніми характеристиками в багатовимірному просторі ознак (*Liu et al., 2022*).

Реалізація цифрової моделі вимагає використання універсальних програмних засобів, здатних ефективно працювати з великими масивами геолого-геофізичних даних. Такий інструментарій забезпечує єдиний підхід до підготовки даних, формування ознак та оцінки якості моделей. Це особливо важливо у випадку складних природних систем, де різномірність інформації потребує її узгодження та структуризації. У сучасних геонаукових дослідженнях поєднання методів аналізу даних із інформаційними технологіями розглядається як необхідна умова побудови адекватних цифрових моделей (*Kempler & Mathews, 2017*).

Таким чином, запропонований у роботі підхід до створення цифрового близнюка передбачає послідовну реалізацію кількох ключових етапів: формування та підготовку вхідного масиву даних, побудову багатовимірному простору ознак, застосування методів класифікації та подальшу перевірку гіпотези подібності. Така структура дозволяє системно організувати процес аналізу та забезпечити логічний зв'язок між окремими етапами дослідження.

Застосування цієї методології дає змогу розглядати цифровий близнюк не лише як інструмент відображення геологічного об'єкта, а як повноцінну аналітичну систему. Вона забезпечує інтеграцію різнорідних даних, їх постійне уточнення та використання для аналізу і прогнозування поведінки пластів-колекторів у процесі експлуатації ПСГ. Це, у свою чергу, створює надійне підґрунтя для прийняття обґрунтованих інженерних рішень.

2.3. Світовий досвід використання цифрових близнюків підземних сховищ газу

Світовий досвід застосування цифрових близнюків у нафтогазовій галузі свідчить про поступовий перехід від традиційних статичних моделей до більш складних динамічних систем, здатних адаптуватися до змін у режимі реального часу. Особливо актуальним це є для підземних сховищ газу, оскільки їх експлуатація характеризується циклічністю технологічних процесів, що супроводжується суттєвими змінами пластового тиску, температури та фазового стану флюїдів.

У таких умовах класичні підходи до моделювання не завжди дозволяють адекватно відобразити реальний стан об'єкта, тоді як цифровий близнюк забезпечує можливість інтеграції геолого-геофізичних та експлуатаційних даних у єдине інформаційне середовище та їх постійного оновлення.

У сучасних дослідженнях цифровий близнюк розглядається як інтегрована система, що поєднує фізичний об'єкт, його цифрову модель та потоки даних між ними. Такий підхід передбачає створення єдиного інформаційного середовища, у якому дані використовуються для постійного оновлення моделі та аналізу стану системи. У випадку підземних резервуарів це дозволяє інтегрувати геолого-геофізичні та експлуатаційні дані для оцінки і прогнозування поведінки об'єкта (*Tao et al., 2019*).

Практична ефективність цифрових близнюків для підземних сховищ газу підтверджується результатами конкретних досліджень. Зокрема, у роботі Елії Абдо розглянуто застосування цифрового близнюка для виявлення витоків газу в свердловинах. Модель базується на аналізі температурних і тискових параметрів, що дозволяє не лише фіксувати відхилення від нормального режиму роботи, але й оцінювати масштаби можливих пошкоджень.

Важливим є те, що цифровий близнюк у цьому випадку виконує не тільки діагностичну, а й прогностичну функцію, дозволяючи завчасно виявляти потенційно небезпечні ситуації та підвищувати рівень безпеки експлуатації. (*Abdo et al., 2023*).

Окрім задач забезпечення технічної цілісності, цифрові близнюки активно застосовуються для оптимізації роботи свердловин і управління режимами експлуатації підземних резервуарів. У дослідженні Н. Атібодхі показано, що інтеграція даних з великої кількості свердловин із математичними моделями дозволяє оперативно коригувати виробничі параметри залежно від фактичного стану системи. Такий підхід дає змогу підвищити ефективність роботи об'єкта, а також мінімізувати ризики, пов'язані з нестабільністю режимів експлуатації. Важливо, що подібні системи функціонують у режимі, близькому до реального часу, що суттєво розширює можливості управління. (*Atibodhi et al., 2023*).

Для більш повного відтворення процесів, які відбуваються в підземних резервуарах, необхідно враховувати їх багатомасштабний характер. У роботі Чжан Тао, запропоновано підхід до побудови цифрового близнюка, який поєднує процеси різного рівня — від мікроскопічних явищ у поровому просторі до макроскопічних гідродинамічних процесів у пласті. Такий підхід дозволяє більш точно описувати поведінку резервуара, оскільки враховує складну взаємодію фізичних факторів. Для підземних сховищ газу це є особливо важливим, адже їх функціонування також визначається взаємодією процесів на різних рівнях. (*Zhang et al., 2020*).

Подібні підходи використовуються не лише для природного газу, а й для інших типів підземного зберігання, зокрема водню та CO₂. У дослідженні Абхінава Пракаша Ґалота цифрові моделі застосовуються для аналізу процесів закачування CO₂ у підземні резервуари з урахуванням геомеханічних факторів. Отримані результати свідчать про можливість визначення безпечних режимів експлуатації та прогнозування змін властивостей порід.

У свою чергу, дослідження у сфері підземного зберігання водню демонструють, що цифрові моделі можуть використовуватися для оцінки втрат газу та аналізу впливу циклічних навантажень на стан резервуара. При цьому варто зазначити, що роботи, присвячені водневому зберіганню, підкреслюють складність таких систем і необхідність врахування геохімічних та геомеханічних процесів, що додатково обґрунтовує застосування комплексних цифрових моделей. (*Gahlot et al., 2024; Yilmaz, 2024*).

Окрему роль у сучасних цифрових близнюках відіграють методи машинного навчання, які дозволяють ефективно працювати з великими обсягами геолого-геофізичних даних. Використання таких методів дає можливість виявляти приховані закономірності, виконувати класифікацію об'єктів та підвищувати точність прогнозування.

Зокрема, у роботах Сат'я Нараян, показано ефективність алгоритмів машинного навчання для класифікації літофацій на основі свердловинних і сейсмічних даних, тоді як Вей Лю демонструють можливості застосування таких підходів для ідентифікації резервуарів і прогнозування їхніх характеристик. Для даного дослідження це є особливо важливим, оскільки підтверджує доцільність використання непараметричних методів класифікації для аналізу подібності геофізичних характеристик пластів (*Narayan et al., 2023; Liu et al. 2022*).

Узагальнюючи світовий досвід, можна зробити висновок, що цифрові близнюки у сфері підземного зберігання газу розвиваються у кількох основних напрямках:

- інтеграція геофізичних і експлуатаційних даних;
- побудова динамічних моделей резервуарів;
- використання алгоритмів машинного навчання;
- прогнозування поведінки системи;
- підтримка прийняття інженерних рішень.

Таким чином, цифровий близнюк у сучасній практиці розглядається як комплексна система, що поєднує дані, моделі та аналітичні методи. Для підземних сховищ газу це означає можливість переходу до нового рівня управління, заснованого на аналізі даних і прогнозуванні.

Висновки до розділу: Підбиваючи підсумок другого розділу, можна стверджувати, що концепція цифрового близнюка є закономірним етапом розвитку методів моделювання у геофізиці, який забезпечує перехід від статичних моделей до динамічних систем, здатних відображати змінний стан підземних резервуарів. Встановлено, що ефективність таких моделей значною мірою залежить від якості вхідних геолого-геофізичних даних та їх представлення у вигляді багатовимірного простору ознак.

Аналіз підходів показав доцільність застосування методів машинного навчання для роботи з неоднорідними геологічними параметрами, що дозволяє формалізувати процес їх аналізу. Водночас узагальнення світового досвіду підтвердило ефективність цифрових близнюків у задачах моніторингу та прогнозування, але також виявило обмеженість підходів до оцінки подібності геофізичних характеристик.

РОЗДІЛ 3. МЕТОДОЛОГІЯ НЕПАРАМЕТРИЧНОЇ КЛАСИФІКАЦІЇ ПРИ ПОБУДОВІ МОДЕЛІ СХОВИЩА

3.1 Теоретичні основи кластер-аналізу та самоорганізовуваних мереж Кохонена

Актуальний етап розвитку геонаук характеризується активним впровадженням методів інтелектуального аналізу багатовимірних даних. Це пов'язано зі значним збільшенням обсягів геолого-геофізичної інформації, яка формується під час сейсморозвідки, каротажу, петрофізичних досліджень та моніторингу пластових систем.

Оскільки петрофізичним характеристикам властива природна неоднорідність і складний нелінійний характер взаємозв'язків, застосування традиційних статистичних підходів часто виявляється недостатньо ефективним. Саме тому сьогодні дедалі ширше використовуються методи машинного навчання та непараметричної класифікації, які дозволяють виявляти приховані закономірності у великих масивах даних (Kohonen, 2001).

Провідним інструментом непараметричного структурування даних у цьому напрямі є кластерний аналіз. Його фундаментальна ідея полягає в автоматичному об'єднанні спостережень у гомогенні групи (кластери) на основі їхньої внутрішньої подібності без використання заздалегідь визначених класів.

Завдяки здатності працювати зі складними багатопараметричними масивами, алгоритми кластеризації дозволяють виявляти природні угруповання навіть у випадках, коли межі між окремими класами є нечіткими або поступовими (Sagir, Sathasivam, 2016).

У контексті аналізу об'єктів підземного зберігання газу такий підхід набуває особливого значення, оскільки геолого-геофізичні дані характеризуються високою багатовимірністю та складними нелінійними взаємозв'язками між параметрами. У

процесі інтерпретації одночасно враховується велика кількість атрибутів, тому звичайне порівняння окремих каротажних або сейсмічних характеристик не завжди дозволяє виявити реальну подібність геологічних середовищ.

Саме тому кластерний аналіз дає можливість комплексно аналізувати сукупність ознак та оцінювати ступінь подібності пластів-колекторів на основі їхніх петрофізичних характеристик (*Smith, Treitel, 2011*).

Одним із найбільш ефективних інструментів реалізації такого підходу є самоорганізовані карти Кохонена (Self-Organizing Maps, SOM). Запропонований фінським дослідником Теуво Кохоненом, цей метод став одним із базових інструментів аналізу багатовимірних систем.

Основна особливість SOM полягає у здатності перетворювати багатовимірні дані у двовимірну карту зі збереженням топологічних зв'язків між об'єктами. Це дозволяє не лише автоматично виконувати кластеризацію, а й візуально аналізувати структуру взаємозв'язків між сформованими групами даних (*Kohonen, 2001*).

Базовим елементом структури SOM є нейрон, який описується власним вектором вагових коефіцієнтів. Під час навчання мережа аналізує вхідні дані та поступово змінює значення ваг таким чином, щоб подібні спостереження групувалися у сусідніх комірках вихідної сітки.

У результаті формується двовимірна карта, де розташування кластерів відображає ступінь подібності між об'єктами. Саме властивість топологічного збереження зв'язків робить SOM ефективним інструментом для аналізу складних багатовимірних даних (*Peeters et al., 2006*).

На противагу класичним нейромережам прямого поширення, карти Кохонена здатні кластеризувати масиви інформації автономно, без попереднього задання цільових класів. В основі їхньої роботи лежить принцип конкурентного навчання. Під час подачі багатовимірних даних мережа визначає нейрон, який

найточніше відповідає параметрам вхідного сигналу. Такий елемент набуває статусу найбільш відповідного вузла (Best Matching Unit, BMU).

Подальший етап навчання полягає у коригуванні вагових коефіцієнтів як самого BMU, так і його сусідніх нейронів, що дозволяє зберігати топологічну структуру даних та формувати карту подібності між об'єктами (*Kohonen, 2001*).

Ефективність методології SOM у науках про Землю підтверджується її широким застосуванням для аналізу сейсмічних атрибутів та класифікації літофаціальних комплексів.

Завдяки здатності структурувати складну багатопараметричну інформацію, алгоритм Кохонена дозволяє виявляти приховані структурні особливості геологічного середовища та просторові закономірності у хвильових полях (*Du et al., 2015*).

Важливою перевагою SOM є їхня здатність працювати з даними, які мають складну нелінійну структуру. У реальних системах залежності між параметрами рідко бувають лінійними, а межі між окремими групами даних часто є нечіткими. Саме тому традиційні методи класифікації можуть давати нестабільні результати. SOM, навпаки, дозволяють враховувати нелінійні взаємозв'язки між параметрами та формувати більш гнучкі моделі кластеризації (*Cottrell et al., 2018*).

Однією з ключових особливостей мереж Кохонена є можливість наочного відображення багатопараметричних даних у вигляді двовимірної карти. У контексті інтерпретації геоданих це відкриває додаткові можливості для аналізу структури даних. Дослідник отримує змогу не лише виконати автоматичну класифікацію, а й візуально оцінити взаємне розташування кластерів, ступінь їхньої відокремленості та характер переходів між суміжними групами.

Оскільки реальні пластові системи характеризуються високою просторовою неоднорідністю, такий підхід дозволяє значно точніше інтерпретувати результати структурування геофізичної інформації (*Baço et al., 2005*).

Отже, кластер-аналіз та самоорганізовувані мережі Кохонена є ефективними інструментами для роботи зі складними багатовимірними геолого-геофізичними даними. Їх використання дозволяє не лише автоматизувати процес класифікації, а й виявляти приховані закономірності, оцінювати ступінь подібності між об'єктами та візуалізувати структуру даних у зручному для інтерпретації вигляді.

Завдяки здатності працювати з нелінійними взаємозв'язками та просторово неоднорідними системами, SOM є особливо перспективними для аналізу пластів-колекторів і створення цифрових близнюків підземних сховищ газу.

3.2 Алгоритм підготовки та структурування багатопараметричних масивів даних для ПСГ

Етап попередньої підготовки та структурування вихідних даних є фундаментальною передумовою для коректного застосування алгоритмів машинного навчання в геонауках. У контексті дослідження об'єктів підземного зберігання газу ця задача постає особливо гостро. Комплексна геолого-петрофізична інформація щодо таких структур зазвичай синтезується з розрізнених джерел, що неминуче зумовлює структурну неузгодженість, різний ступінь деталізації вимірів та наявність пропусків.

З огляду на це, формування зведеного багатопараметричного масиву для класифікації виходить за межі простої компіляції показників; воно вимагає ретельного попереднього аналізу, стандартизації та цілеспрямованої адаптації матеріалу для подальшого алгоритмічного опрацювання (*Karpatne et al., 2017*).

Впровадження підходів, керованих даними, та методів машинного навчання у сучасну геонауку безпосередньо спирається на створення багатопараметричних баз даних. Їхня побудова вимагає об'єднання розрізненої геологічної, геофізичної та петрофізичної інформації в єдиний інформаційний простір (*Bergen et al., 2019*).

З практичної точки зору для задач класифікації найбільш придатною є таблична структура даних, у якій об'єкти аналізу розташовуються по рядках, а їхні ознаки — по стовпцях. При дослідженні підземних сховищ газу ключовим об'єктом класифікації переважно обирають конкретний пласт у межах окремої свердловини. Такий підхід є обґрунтованим, оскільки дозволяє інтегрувати стратиграфічні, літофаціальні та геофізичні характеристики на єдиному рівні аналізу.

Формування аналітичної бази розпочинається з аудиту доступних геологічних матеріалів. Для структур підземних сховищ газу робочий датасет зазвичай інтегрує матеріали свердловинної геофізики, цифрові каротажні криві, стратиграфічні прив'язки, петрофізичні властивості за керном та результати попередніх етапів інтерпретації.

Найбільшу аналітичну цінність при цьому становлять дані ГДС, які дають змогу кількісно охарактеризувати розріз за показниками гамма-активності, електричного опору, густини, акустичних характеристик та пористості. У сучасних підходах до автоматизованої геофізичної інтерпретації саме комплекс well log параметрів розглядається як базова основа для вирішення задач фаціальної класифікації методами машинного навчання (*Bestagini et al., 2017*).

Специфіка опрацювання свердловинних матеріалів значною мірою зумовлена їхньою структурною неоднорідністю, що полягає у поєднанні безперервних кривих ГДС із категоріальними даними. У роботах, присвячених machine learning для facies classification, каротажні параметри розглядаються як система взаємопов'язаних предикторів, на основі яких формується навчальна вибірка (*Hall, 2016*).

Саме тому одним із обов'язкових етапів preprocessing є уніфікація структури даних — приведення різноформатної інформації до єдиного стандарту з однаковим набором колонок для кожного об'єкта аналізу. Типова структура такого масиву зазвичай включає просторові параметри («Свердловина», «Пласт», глибина покрівлі та підшви, товщина), набір стандартизованих петрофізичних

характеристик (GR, Resistivity, Density, Sonic, Porosity), а також категоріальні ознаки («Літологія», «Тип колектора»).

Таблиця 3.1. Структура багатопараметричного набору даних для ML-класифікації.

(Джерело: Hall, 2016)

	Facies	Depth	GR	ILD_ log10	DeltaPHI	PHIND	PE	NM_M	RELPOS
count	3232	3232	3232	3232	3232	3232	3232	3232	3232
mean	4.42	2875.82	66.14	0.64	3.55	13.48	3.73	1.50	0.52
std	2.50	131.00	30.85	0.24	5.23	7.70	0.89	0.50	0.29
min	1	2573.50	13.25	-0.03	-21.83	0.55	.020	1	0.01
25%	2	2791.00	46.92	0.49	1.16	8.35	3.10	1	0.27
50%	4	2932.50	65.72	0.62	3.5	12.15	3.55	2	0.53
75%	6	2980.00	79.63	0.81	6.43	16.45	4.30	2	0.77
max	9	3122.50	361.15	1.48	18.60	84.40	8.09	2	1.00

Оскільки кожен досліджуваний інтервал репрезентований не ізольованими значеннями, а багатовимірною системою взаємопов'язаних ознак із прихованими нелінійними залежностями, першочерговим завданням є попередня статистична оцінка діапазонів варіативності та структурної неоднорідності вибірки. Саме такий аналіз забезпечує коректну підготовку геоданих до подальшої ML-класифікації та створює математичне підґрунтя для застосування алгоритмів кластер-аналізу й самоорганізовуваних карт Кохонена (SOM).

Водночас ефективність подальшого моделювання значною мірою залежить від якості вихідного набору даних, тому наступним важливим етапом є їхня попередня перевірка та уніфікація.

Для геолого-геофізичних баз характерними проблемами є дублікати, технічні похибки, пропуски та різні одиниці вимірювання. Наприклад, пористість може бути представлена як у відсотках, так і в частках одиниці, а глибина — у метрах або кілометрах.

Якщо заздалегідь не усунути таку неоднорідність, алгоритм почне некоректно оцінювати співвідношення між параметрами, що безпосередньо

вплине на результати класифікації. У фаховій літературі з preprocessing наголошується, що ефективне застосування алгоритмів Data Mining можливе лише після очищення, інтеграції та трансформації вихідних даних у стандартизований формат (*García et al., 2015*).

Однією з найбільш складних проблем preprocessing є обробка пропущених значень. У свердловинних даних пропуски можуть виникати через відсутність окремих методів каротажу, втрату частини кривих або неповну цифрову обробку архівних матеріалів. У сучасних дослідженнях із machine learning для geoscience підкреслюється, що некоректна робота з missing values здатна призводити до формування штучних закономірностей у моделі та суттєвого зниження її прогностичної здатності (*Veerman et al., 2021*).

Практична робота з пропущеними значеннями може реалізовуватися декількома способами. Якщо кількість пропусків є незначною, окремі записи можуть бути вилучені з робочого масиву. Для високоінформативних параметрів допускається використання статистичних методів імпутації, зокрема підстановка середніх або медіанних значень у межах певного геологічного інтервалу.

Водночас надмірне штучне відновлення каротажних параметрів є небажаним, оскільки воно може спотворювати природну структуру геологічних взаємозв'язків. У роботах, присвячених machine learning для facies classification, підкреслюється, що ефективність автоматизованої інтерпретації безпосередньо залежить від якості вхідних ознак, preprocessing та feature engineering (*Guarido, 2019*).

Одним із найважливіших етапів є узгодження даних за глибиною та стратиграфією. Каротажні криві характеризуються високою дискретністю запису, тоді як стратиграфічні таблиці містять лише межі пластів або горизонтів. Для інтеграції цих даних необхідно перейти від безперервних кривих ГДС до дискретних пластових параметрів.

З цією метою в межах кожного інтервалу розраховуються середні, медіанні, мінімальні та максимальні значення каротажних показників. У сучасних підходах до *facies classification well log data* розглядаються саме як система взаємопов'язаних ознак, придатних для подальшого застосування алгоритмів машинного навчання (*Bestagini et al., 2017*).

Важливою складовою *preprocessing* є формування похідних ознак. Первинними параметрами вважаються значення, безпосередньо отримані з каротажних або лабораторних досліджень. Похідні ознаки формуються шляхом математичної обробки первинних параметрів.

Наприклад, товщина пласта визначається як різниця між глибиною підошви та покрівлі, а середні значення каротажних параметрів можуть розраховуватися в межах окремого інтервалу. У задачах фаціальної класифікації додаткове конструювання ознак на основі *well log* параметрів дозволяє підвищити інформативність навчального масиву та покращити якість подальшого розпізнавання геологічних класів (*Guarido, 2019*).

Наступним етапом є кодування категоріальних параметрів. Для геолого-геофізичних задач такими ознаками можуть бути літологія, стратиграфічний вік, тип колектора або характер флюїдонасичення. У процесі цифрової обробки ці параметри необхідно переводити у формат, придатний для роботи алгоритмів машинного навчання.

Для цього використовуються підходи, пов'язані з трансформацією категоріальних змінних у машинно-читаний формат. Водночас важливо уникати створення хибної рангової залежності між категоріями, якщо така залежність геологічно не обґрунтована (*García et al., 2015*).

Після завершення *preprocessing* формується фінальний робочий масив, який поділяється на матрицю ознак та цільову змінну. Матриця ознак містить усі геолого-геофізичні параметри, необхідні алгоритму для побудови дерева рішень, а цільова змінна визначає клас або параметр, що підлягає прогнозуванню.

У задачах аналізу ПСГ цільовою змінною можуть виступати тип літології, клас колекторських властивостей або ступінь подібності геологічних об'єктів між собою. Загалом такий підхід відповідає сучасній концепції інтегрованої характеристики резервуарів, у межах якої поєднання геологічних знань і вимірних даних використовується для побудови більш обґрунтованих моделей підземного середовища (*Ma & La Pointe, 2011*).

Наступним етапом preprocessing є розподіл даних на навчальну та тестову вибірки. Навчальний масив використовується для побудови класифікаційної моделі, тоді як тестова вибірка застосовується для перевірки її прогностичної здатності. Для геолого-геофізичних задач особливо важливо контролювати не лише статистичну, а й геологічну репрезентативність вибірок.

Якщо в навчальному масиві переважають лише певні типи пластів або літологічних умов, модель може демонструвати високу формальну точність під час навчання, але втрачати ефективність у нових геологічних ситуаціях (*Karpatne et al., 2017*).

Отже, підготовка та структурування багатопараметричних масивів даних є одним із центральних етапів застосування методу Кохонена у задачах аналізу ПСГ. Саме на цьому етапі різномірна геологічна, геофізична та петрофізична інформація перетворюється на впорядковану систему ознак, придатну для подальшого машинного аналізу.

Якісно сформований масив даних забезпечує не лише високу точність класифікації, а й збереження геологічного змісту отриманих результатів. Сучасний підхід до характеристики резервуарів передбачає поєднання алгоритмічного аналізу з повноцінною геологічною інтерпретацією середовища (*Ma & La Pointe, 2011*).

3.3. Процедури крос-перевірки та верифікації прогностичної здатності моделі

Одним із ключових етапів застосування алгоритмів машинного навчання є перевірка їхньої здатності коректно працювати не лише з навчальними даними, а й з новими вибірками, які не використовувалися під час побудови моделі. Для задач геолого-геофізичної інтерпретації це питання має особливе значення, оскільки навіть високі показники формальної точності моделі не завжди гарантують її геологічну достовірність.

Саме тому в сучасних дослідженнях значна увага приділяється процедурам крос-перевірки (cross-validation) та верифікації прогностичної здатності алгоритмів машинного навчання (*Hastie et al., 2009*).

У загальному розумінні верифікація моделі полягає у перевірці того, наскільки коректно алгоритм здатний узагальнювати закономірності, виявлені в навчальному масиві даних. Для алгоритмів непараметричної кластеризації проблема перенавчання є особливо актуальною, оскільки дерево рішень може надмірно адаптуватися до особливостей тренувальної вибірки.

У такому випадку модель починає враховувати випадкові особливості та локальні аномалії замість реальних закономірностей. У результаті дерево демонструє високу точність на навчальних даних, але втрачає ефективність під час роботи з новими геологічними об'єктами (*Breiman et al., 1984*).

Проблема перенавчання є однією з центральних у machine learning. У геолого-геофізичних задачах вона додатково ускладнюється високою неоднорідністю природних систем. Навіть у межах одного родовища петрофізичні характеристики можуть суттєво змінюватися залежно від літології, фаціальних умов або ступеня флюїдонасичення.

Саме тому модель, яка демонструє високу точність у межах однієї вибірки, не завжди здатна адекватно працювати в інших геологічних умовах. У сучасних

geoscience-дослідженнях наголошується, що перевірка узагальнювальної здатності моделі є критично важливою умовою її практичного використання (*Karpatne et al., 2017*).

Найпростішим способом оцінки якості моделі є розподіл даних на навчальну та тестову вибірки. У цьому випадку одна частина масиву використовується для побудови дерева рішень, а інша — для перевірки його роботи. Найчастіше використовуються співвідношення 70/30 або 80/20, де більша частина даних застосовується для навчання, а менша — для тестування. Такий підхід дозволяє оцінити, наскільки стабільно модель працює на даних, які не були використані під час її побудови (*Kuhn & Johnson, 2013*).

Проте простий train-test split має певні обмеження. Результат оцінки значною мірою залежить від того, які саме об'єкти потрапили до навчальної та тестової вибірок. Якщо тестовий масив випадково виявиться простішим або, навпаки, складнішим за тренувальний, результати можуть бути викривлені. Саме тому в сучасних дослідженнях перевага надається більш стійким методам перевірки, насамперед крос-валідації (*James et al., 2021*).

Крос-перевірка (cross-validation) — це процедура багаторазового тестування моделі на різних підмножинах даних. Її основна ідея полягає в тому, що кожен об'єкт вибірки певний час виступає як елемент навчального масиву, а певний — як елемент тестового. Завдяки цьому оцінка якості моделі стає менш залежною від випадкового поділу даних та більш об'єктивно відображає її прогностичну здатність (*Hastie et al., 2009*).

Найбільш поширеним методом є k-fold cross-validation (Див. Рис.3.2). У межах цього підходу весь масив даних поділяється на k рівних частин (folds). На кожній ітерації одна частина використовується для тестування, а решта — для навчання моделі.

Процедура повторюється доти, поки кожен fold не буде використаний як тестовий. Після завершення всіх ітерацій обчислюється середнє значення

показників точності, що дозволяє отримати більш стабільну оцінку якості моделі (Kuhn & Johnson, 2013).

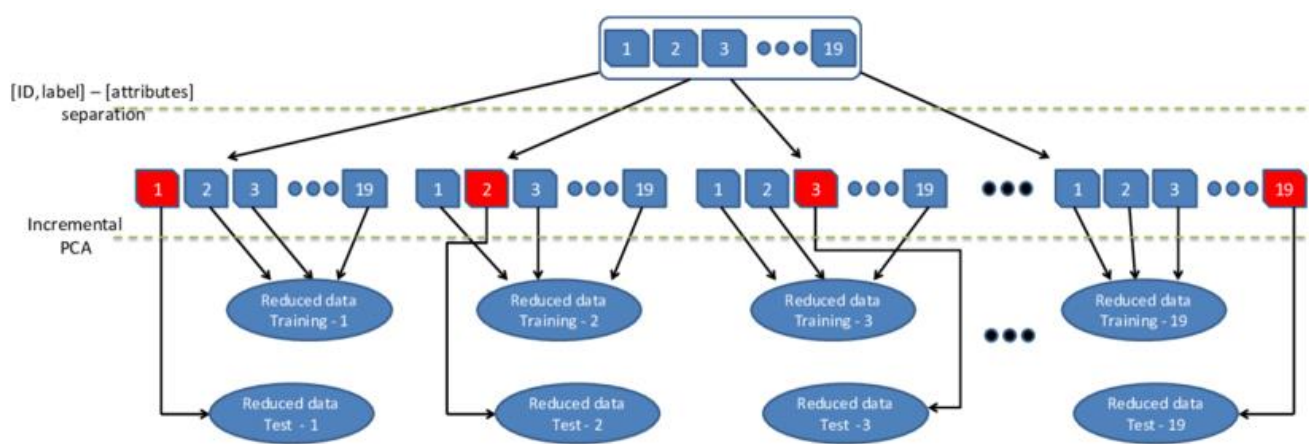


Рис.3.2 - Схема k-fold крос-валідації

(Джерело: Dehghani & Azadifar, 2016)

На рисунку представлено схему k-fold крос-валідації — одного з найбільш поширених методів перевірки якості моделей машинного навчання. У верхній частині схеми показано початковий набір даних, що складається з ідентифікатора об'єкта, класу та набору атрибутів ([ID, label] – [attributes]). Далі масив інформації послідовно розділяється на окремі підвибірки (folds), де одна частина використовується як тестова вибірка, а інші — для навчання моделі.

На схемі також відображено етап поетапного зменшення розмірності даних за допомогою методу головних компонент (Incremental PCA), після чого формуються навчальні (Reduced data Training) та тестові (Reduced data Test) набори даних для кожної ітерації перевірки. Такий підхід дозволяє мінімізувати вплив випадкового розподілу вибірки, оцінити стійкість алгоритму до змін вхідних даних та підвищити надійність результатів класифікації.

Для геолого-геофізичних задач це має особливе значення, оскільки багатопараметричні набори геоданих часто характеризуються високою неоднорідністю та складними нелінійними взаємозв'язками між параметрами.

Саме тому важливим є правильний вибір кількості підвбірок (folds), від яких залежить стабільність та достовірність оцінки моделі.

Для задач класифікації найчастіше використовується 5-fold або 10-fold cross-validation. Вважається, що саме такі значення забезпечують оптимальний баланс між точністю оцінки та обчислювальними витратами. У геолого-геофізичних задачах 10-fold cross-validation використовується особливо часто, оскільки дозволяє більш стабільно оцінювати роботу алгоритму на неоднорідних наборах даних. дозволяє ефективно працювати навіть із відносно невеликими вибірками свердловинних даних (*James et al., 2021*).

Окремим випадком крос-перевірки є leave-one-out cross-validation (LOOCV). У цьому методі кожен об'єкт по черзі використовується як тестовий, тоді як усі інші формують навчальну вибірку. Такий підхід забезпечує максимально повне використання доступних даних, проте потребує значних обчислювальних ресурсів. Для великих геолого-геофізичних масивів він використовується рідше, однак може бути корисним при роботі з обмеженою кількістю свердловин або пластів (*Hastie et al., 2009*).

Важливим елементом верифікації моделі є оцінка метрик класифікації. Найпростішим показником є ассигасу — частка правильно класифікованих об'єктів. Проте для геолого-геофізичних задач використання лише ассигасу часто є недостатнім. Якщо певні класи представлені нерівномірно, модель може демонструвати високу загальну точність навіть за умови поганої класифікації рідкісних, але геологічно важливих об'єктів (*Sokolova & Lapalme, 2009*).

Саме тому додатково використовуються precision, recall та F1-score. Precision показує, яка частина об'єктів, віднесених моделлю до певного класу, дійсно належить до нього. Recall характеризує здатність алгоритму виявляти всі об'єкти певного класу. F1-score є гармонійним середнім між precision та recall і дозволяє комплексно оцінити якість класифікації (*Sokolova & Lapalme, 2009*).

Для наочного аналізу результатів класифікації широко використовується confusion matrix — матриця помилок. Вона відображає співвідношення між реальними та прогнозованими класами й дозволяє визначити, які саме типи об'єктів модель плутає між собою. У геофізичних задачах це має особливе значення, оскільки допомагає оцінити, наскільки добре алгоритм розділяє близькі за властивостями літологічні або колекторські типи (*Provost & Fawcett, 2013*).

Оскільки структура дерева формується на основі конкретної вибірки, навіть незначні зміни у вхідних даних можуть призводити до зміни ієрархії розгалужень. Саме тому рекомендується оцінювати не лише загальні метрики точності, а й стійкість моделі до варіацій вхідного масиву (*Breiman et al., 1984*).

На етапі верифікації особливого значення набуває перевірка того, наскільки результати класифікації узгоджуються з реальними геологічними умовами досліджуваного об'єкта. Для геонаук недостатньо лише формальної статистичної точності. Важливо, щоб закономірності, виявлені алгоритмом, мали фізико-геологічне пояснення та відповідали реальним властивостям середовища.

У сучасних geoscience-дослідженнях підкреслюється, що моделі машинного навчання повинні узгоджуватися з фізичними процесами та геологічною будовою об'єкта дослідження (*Bergen et al., 2019*).

Під час верифікації моделей важливо не лише оцінити статистичні метрики, а й проаналізувати логіку прийняття рішень алгоритмом. Це спрощується завдяки прозорій структурі дерева рішень, яка дозволяє простежити послідовність умов класифікації та визначити, які саме параметри найбільше впливають на результат моделювання. У сучасних дослідженнях такий підхід розглядається як складова концепції explainable AI — інтерпретованого машинного навчання (*Molnar, 2022*).

Підсумовуючи, перехресна перевірка та експертна верифікація виступають методологічним базисом застосування алгоритму КохоненаС. Вони забезпечують перехід від формальної оцінки точності до підтвердження предиктивного потенціалу та стабільності моделі.

Для завдань моделювання підземних сховищ газу цей етап є визначальним, адже надійність результатів вимагає суворої відповідності згенерованих математичних правил фактичній будові геологічного середовища.

Висновки до розділу: Фундаментальною базою для розробки цифрового близнюка ПСГ є комплексний аналіз багатовимірних геолого-геофізичних масивів. Пріоритетними інструментами тут виступають алгоритми кластеризації та самоорганізовувані карти Кохонена (SOM), які ефективно опрацьовують складну архітектуру геоданих і фіксують приховані закономірності.

Водночас точність фінальної класифікації та об'єктивність тестування гіпотези про подібність об'єктів безпосередньо залежать від попередньої статистичної підготовки вибірки. Саме якість вхідної інформації є головною запорукою достовірності майбутньої цифрової моделі.

РОЗДІЛ 4. ПРАКТИЧНА РЕАЛІЗАЦІЯ ЦИФРОВОГО БЛИЗНЮКА ТА ПЕРЕВІРКА ГІПОТЕЗИ ПОДІБНОСТІ ГЕОФІЗИЧНИХ ХАРАКТЕРИСТИК

4.1 Формування та підготовка геолого-геофізичної бази даних для моделювання

Практична реалізація дослідження розпочиналася з збору та систематизації вхідних геолого-геофізичних даних, необхідних для подальшого застосування методу. На цьому етапі було сформовано єдину інформаційну базу по Полтавському ПСГ та ряду перуанських родовищ-аналогів, які використовувалися для перевірки гіпотези подібності геофізичних характеристик.

Основна увага приділялася відбору найбільш інформативних параметрів, формуванню зведених таблиць та приведенню різнорідних даних до єдиної структури для подальшого аналізу.

Для формування вхідної бази геолого-геофізичних даних по Полтавському ПСГ мною були використані матеріали, отримані під час проходження виробничої практики в АТ «Укртрансгаз». Основна частина даних була представлена у вигляді геологічної документації, стратиграфічних матеріалів, структурних побудов, карт пористості та геофізичних каротажних даних свердловин (Див. Рис.4.1).

На основі цих матеріалів у подальшому формувалася єдина зведена таблиця, яка використовувалася для побудови моделі непараметричної кластеризації методами Кохонена та кластерного аналізу.

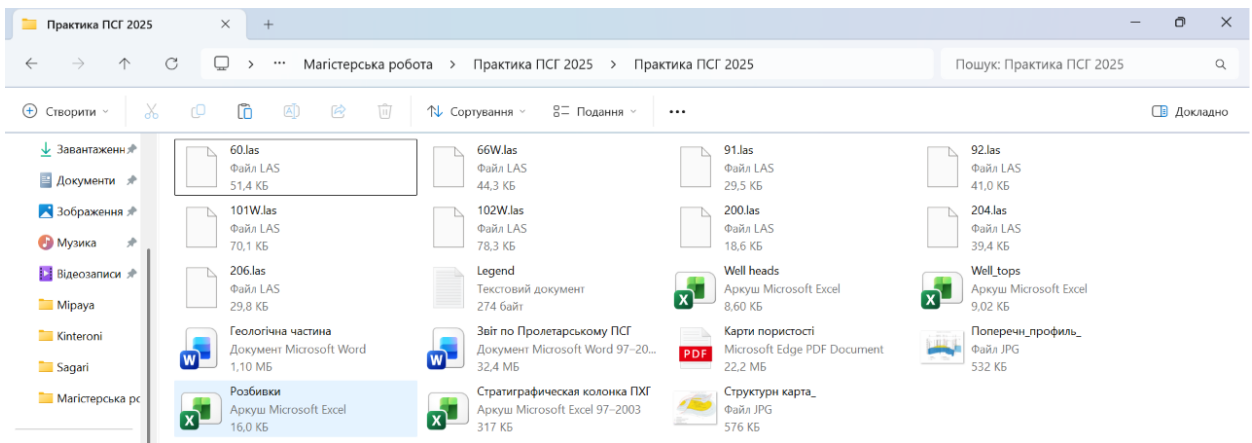


Рис.4.1 – Вихідні геолого-геофізичні матеріали по Полтавському ПСГ

У роботі були використані таблиці Well heads та Well tops, які містили інформацію про свердловини, стратиграфічні межі пластів та глибини залягання горизонтів. Також використовувалася стратиграфічна колонка Полтавського ПСГ, що дозволила виконати прив'язку пластів до відповідних геологічних систем, ярусів та продуктивних комплексів.

Документ «Геологічна частина» містив опис геологічної будови району, літологічних особливостей, стратиграфії та історії геологічного вивчення структури. Окремо були використані структурні карти та поперечні геологічні профілі, які дозволили оцінити просторову будову Полтавського ПСГ, положення продуктивних горизонтів та характер тектонічних порушень.

Важливу роль у формуванні бази даних відіграли LAS-файли свердловин, отримані по ряду експлуатаційних та розвідувальних свердловин Полтавського ПСГ. Саме з цих файлів були отримані основні геофізичні параметри, які надалі використовувалися для формування ознак.

Серед основних параметрів використовувалися дані акустичного каротажу, гамма-каротажу, пористості глинистості, коефіцієнта нафтогазонасичення та інших геофізичних характеристик колекторів. Для різних свердловин набір параметрів частково відрізнявся, що було пов'язано з неоднорідністю вихідних матеріалів та різним ступенем вивченості окремих ділянок структури.

Дані матеріали дозволили виділити зони розвитку колекторів, оцінити просторовий розподіл пористості та визначити найбільш перспективні ділянки в межах структури.

Окремо у роботі використовувався файл «Звіт по Полтавському ПСГ», який містив інформацію щодо ресурсної бази Руденківсько-Пролетарського району, стратиграфічного розподілу ресурсів, інтервалів глибин та характеристик продуктивних комплексів. Частина інформації з даного звіту використовувалася для уточнення стратиграфічної прив'язки пластів та формування окремих параметрів майбутньої зведеної таблиці.

У результаті аналізу та систематизації всіх вищезазначених матеріалів була сформована єдина зведена таблиця по Полтавському ПСГ, у якій по кожному пласту та свердловині були зібрані основні геолого-геофізичні характеристики.

Саме ця таблиця у подальшому стала основою для побудови моделі непараметричної класифікації та перевірки гіпотези подібності геофізичних характеристик між Полтавським ПСГ та перуанськими родовищами-аналогами.

Для формування порівняльної вибірки та подальшої перевірки гіпотези подібності геофізичних характеристик окрім даних по Полтавському ПСГ, у роботі також використовувалися геолого-геофізичні матеріали по ряду газоконденсатних родовищ Перу.

Пошук та підбір вихідних даних здійснювався за допомогою відкритої геологічної плаформи PeruPetro Lumin Explorer (Див Рис.4.2), яка містить значний обсяг геологічної, стратиграфічної та геофізичної інформації по нафтогазових родовищах Перу.

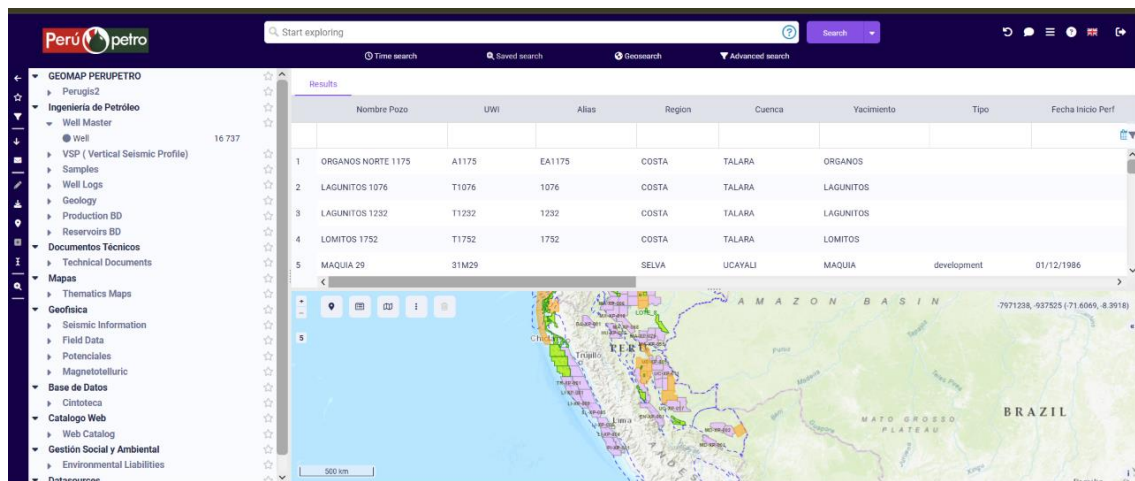


Рис.4.2 - Платформа PeruPetro Lumin Explorer

На відміну від Полтавського ПСГ, де більшість матеріалів була отримана під час проходження виробничої практики, по перуанських родовищах необхідної інформації виконувався окремо для кожного родовища та його свердловин. Основна складність полягала у тому що дані були розміщені у різних вкладках та мали неоднорідну структуру

У процесі роботи мною були проаналізовані різні родовища Перу, серед яких особлива увага приділялася родовищам San Martín, Cashiriari, Pagoreni та Miraya. Вибір саме цих родовищ був пов'язаний із наявністю достатнього обсягу відкритих геофізичних даних, стратиграфічної інформації та каротажних матеріалів, необхідних для формування порівняльної вибірки.

Основний пошук даних виконувався у вкладках Wells, Stratigraphic Units та Well Logs (Див.Рис.4.3). У вкладці Stratigraphic Units здійснювався відбір пластів та стратиграфічних горизонтів для кожної свердловини. Саме ця інформація використовувалася для формування структури майбутніх зведених таблиць та визначення інтервалів продуктивних пластів.

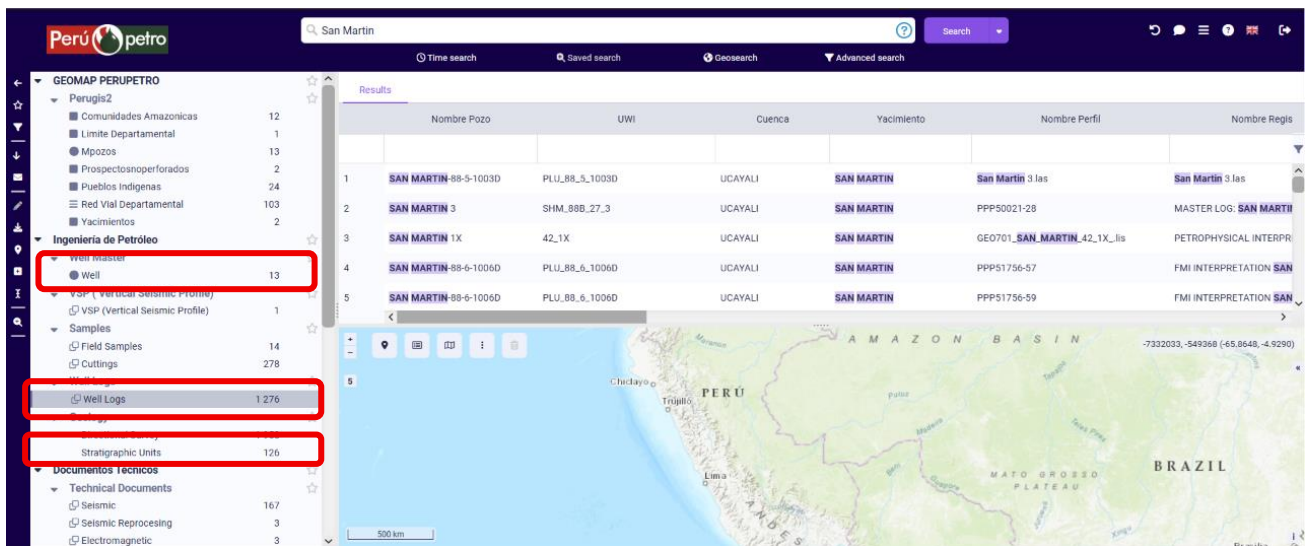


Рис.4.3 - Основні вкладки платформи PeruPetro Lumin Explorer для пошуку геолого-геофізичних даних на прикладі родовища San Martin

Вкладка Well Logs використовувалася для завантаження геофізичних каротажних даних у форматі LAS та TIFF. Зокрема, виконувалося завантаження даних гамма-каротажу, акустичного каротажу, резистивіметрії, щільнісного каротажу та інших геофізичних параметрів, які могли бути використані для подальшого аналізу.

Під час формування вибірки значна увага приділялася підбору спільних геофізичних параметрів між Полтавським ПСГ та Перуанськими родовищами. Оскільки різні свердловини містили неоднакові набори каротажних даних у зведені таблиці включалися лише ті параметри, які повторювалися у більшості досліджуваних об'єктів та могли бути коректно використані для подальшого порівняльного аналізу.

Для кожного родовища окремо формувалися зведені таблиці по пластах та свердловинах (Див рис.4.4), у які заносилися основні геолого-геофізичні характеристики: глибини залягання, товщини пластів, літологічні особливості, пористість, гамма-каротаж, акустичні характеристики, електричний опір та інші параметри, отримані з каротажних матеріалів. У подальшому всі таблиці були

зведені до єдиної структури для можливості їх сумісного аналізу в середовищі Statistica.

Свердловина	Пласт	Глибина мин. м	Глибина макс. м	Інтервал глибин, м	Потужність, м	Гамма-каротаж (GR), API	Електричний опір, Ом-м	Щільність, г/см³	Акустичний каротаж, мкс/м	Пористість, д.од.	Літологія
SAN MARTIN 1X	CUSHABATAY	3863	3894	3000–4000	31	85	8	2,375	90	0,14	пісковики
SAN MARTIN 1X	ESPERANZA	3780	3853	3000–4000	83	80	10	2,45	85	0,115	піщано-глинисті
SAN MARTIN 1X	AGUA CALIENTE	3685	3780	3000–4000	95	75	12,5	2,4	85	0,16	пісковики
SAN MARTIN 1X	CHONTA	3668	3685	3000–4000	17	70	15	2,825	70	0,075	глинисто-карбонатні
SAN MARTIN 1X	AMBO GROUP	3346	3668	3000–4000	322	65	17,5	2,45	85	0,115	пісковики, алеволіти
SAN MARTIN 1X	TARMA GROUP	3223	3346	3000–4000	123	60	21,5	2,675	70	0,075	карбонатно-теригенні
SAN MARTIN 1X	COPACABANA GROUP	2356	3223	2000–3000	867	55	25	2,725	85	0,055	карбонати
SAN MARTIN 1X	ENE	2313	2356	2000–3000	43	50	30	2,5	80	0,125	піщано-глинисті
SAN MARTIN 1X	CUSHABATAY	2222	2313	2000–3000	91	55	22	2,4	85	0,16	пісковики
SAN MARTIN 1X	ESPERANZA	2149	2222	2000–3000	73	50	25	2,45	85	0,115	піщано-глинисті
SAN MARTIN 1X	AGUA CALIENTE	2063	2149	2000–3000	86	47,5	27,5	2,4	85	0,15	пісковики

Рис.4.4 – Фрагмент зведеної таблиці геолого-геофізичних параметрів по родовищу San Martin

Таким чином у результаті проведеної роботи була сформована єдина база геолого-геофізичних даних по Полтавському ПСГ та ряду перуанських родовищ-аналогів, яка стала основою для подальшого застосування методу Кохонена та перевірки гіпотези подібності геофізичних характеристик між досліджуваними об'єктами.

4.2. Побудова моделі непараметричної класифікації ознак колекторів

Для перевірки гіпотези подібності геофізичних характеристик між Полтавським ПСГ та перуанськими родовищами у даній роботі було застосовано методи Кохонена та непараметричної кластеризації. Практична реалізація моделі виконувалася у програмному середовищі Statistica, яке дозволяє проводити багатовимірний аналіз даних, будувати дерева рішень та виконувати класифікацію об'єктів за набором геолого-геофізичних параметрів.

Основою для побудови моделі стали сформовані на попередньому етапі зведені таблиці, які містили геофізичні характеристики пластів та колекторів по

Полтавському ПСГ і перуанських родовищах-аналогах. У процесі роботи було виконано декілька типів аналізу, спрямованих на оцінку подібності досліджуваних об'єктів, визначення найбільш інформативних параметрів та побудову класифікаційних моделей для подальшої верифікації цифрового близнюка.

Для практичної реалізації моделі у даній роботі було використано програмне забезпечення Statistica (Див рис.4.5). Вибір саме цього програмного середовища був пов'язаний із можливістю виконання багатовимірною статистичного аналізу, роботи з великими масивами табличних даних та побудови моделей класифікації на основі дерев рішень. Програма дозволяє поєднувати класичні статистичні методи з елементами машинного навчання, що є особливо важливим при аналізі складних геолого геофізичних параметрів.



Рис.4.5 - Робоче середовище програмного забезпечення Statistica 12

У межах даного дослідження Statistica використовувалася для імпорту сформованих зведених таблиць, аналізу геофізичних характеристик, колекторів та побудови класифікаційних моделей. Перевагою даного програмного забезпечення є можливість візуального представлення дерев рішень, що дозволяє не лише

отримати результат класифікації, а й простежити логіку розподілу об'єктів за окремими геофізичними параметрами.

Ще однією важливою перевагою використання даного програмного середовища стала можливість роботи як із кількісними, так і з якісними параметрами одночасно. Це є особливо актуальним для геологічних задач, де частина ознак має числовий характер (глибини, пористість, електричний опір, щільність), а частина представлена категоріальними характеристиками, такими як літологія чи тип колектора.

У процесі роботи в середовищі Statistica були імпортовані зведені таблиці по Полтавському ПСГ та перуанських родовищах-аналогах, після чого виконувалося формування вибірок для подальшого аналізу. На основі цих даних у подальшому були побудовані декілька моделей класифікації, результати яких використовувалися для оцінки подібності геофізичних характеристик досліджуваних об'єктів.

Після формування зведених таблиць виконувалася підготовка даних до подальшого аналізу в середовищі Statistica. Основним форматом роботи в програмі є табличне представлення даних, де рядки відповідають окремим спостереженням, а стовпці – змінним або геофізичним параметрам (Див.рис.4.6). Такий підхід виявився зручним для роботи з геолого-геофізичним вибірками, оскільки дозволив структуровано організувати дані по пластах та свердловинах.

The screenshot shows a window titled "Data: 2 Nazaruk Артегована (4v by 140с)". It contains a table with 5 columns and 12 rows. The first row is a header, and the subsequent 11 rows represent individual objects. The columns are labeled: "1 objects", "2 Poro, m3/m3", "3 GKK, _", and "4 AK, _".

	1 objects	2 Poro, m3/m3	3 GKK, _	4 AK, _
1	Obj 1	-0,055345	-0,75104	1,59311
2	Obj 2	-0,705018	-0,16636	0,74076
3	Obj 3	-0,752378	0,10700	0,73780
4	Obj 4	-0,722580	-0,06684	0,40686
5	Obj 5	0,124922	-0,77513	0,04690
6	Obj 6	-0,569416	-0,25310	0,03178
7	Obj 7	-0,592240	-0,03094	0,56157
8	Obj 8	-0,718329	-0,38072	1,01092
9	Obj 9	-0,519313	-0,21360	0,64604
10	Obj 10		-0,04504	0,10006
11	Obj 11	0,014000	-0,89988	-0,83711

Рис.4.6 - Імпортована таблиця геолого-геофізичних параметрів у середовищі Statistica

На першому етапі роботи виконувалося створення та імпорт таблиць даних у програмне середовище. Для цього використовувалися сформовані раніше таблиці, які містили основні геофізичні параметри колекторів по Полтавському ПСГ та перуанських родовищах-аналогах. Імпорт даних здійснювався через вбудовані інструменти відкриття Excel- та CSV-файлів, після чого дані автоматично відображалися у вигляді таблиць спостережень і змінних.

У процесі роботи кожен пласт розглядався як окреме спостереження, а основні геолого-геофізичні характеристики використовувалися як змінні моделі. До числа основних параметрів входили:

- Глибина залягання;
- Потужність пласта;
- Гамма-каротаж;
- Електричний опір;
- Щільність;
- Акустичний каротаж;
- Пористість;
- Літологічні характеристики.

Для забезпечення коректної роботи моделі особлива увага приділялася структурі таблиць та типам змінних. Кількісні параметри задавалися у числовому форматі, тоді як окремі літологічні характеристики використовувалися як категоріальні змінні. Такий підхід дозволив одночасно аналізувати як числові, так і якісні характеристики колекторів.

Однією з переваг використання Statistica стала можливість візуального контролю структури даних та подальшого графічного представлення результатів класифікації. Це дозволило не лише виконувати аналіз, а й простежувати логіку побудови дерева рішень та оцінювати вплив окремих геофізичних параметрів на процес класифікації.

Першим етапом практичної реалізації непараметричної класифікації став кластерний аналіз, метою якого була попередня оцінка ступеня подібності геофізичних характеристик між пластами Полтавського ПСГ та перуанських родовищ. Даний тип моделювання ґрунтувався на використанні лише спільних геофізичних ознак, які були присутні у більшості досліджуваних об'єктів.

Необхідність такого підходу була пов'язана з тим, що українська та перуанська вибірки характеризувалися різною кількістю пластів, різною деталізацією геофізичних досліджень та неоднаковими наборами каротажних параметрів. У зв'язку з цим для забезпечення коректного порівняння було сформовано єдиний простір спільних ознак, який включав лише ті параметри, що повторювалися у двох вибірках та могли бути використані для подальшої класифікації.

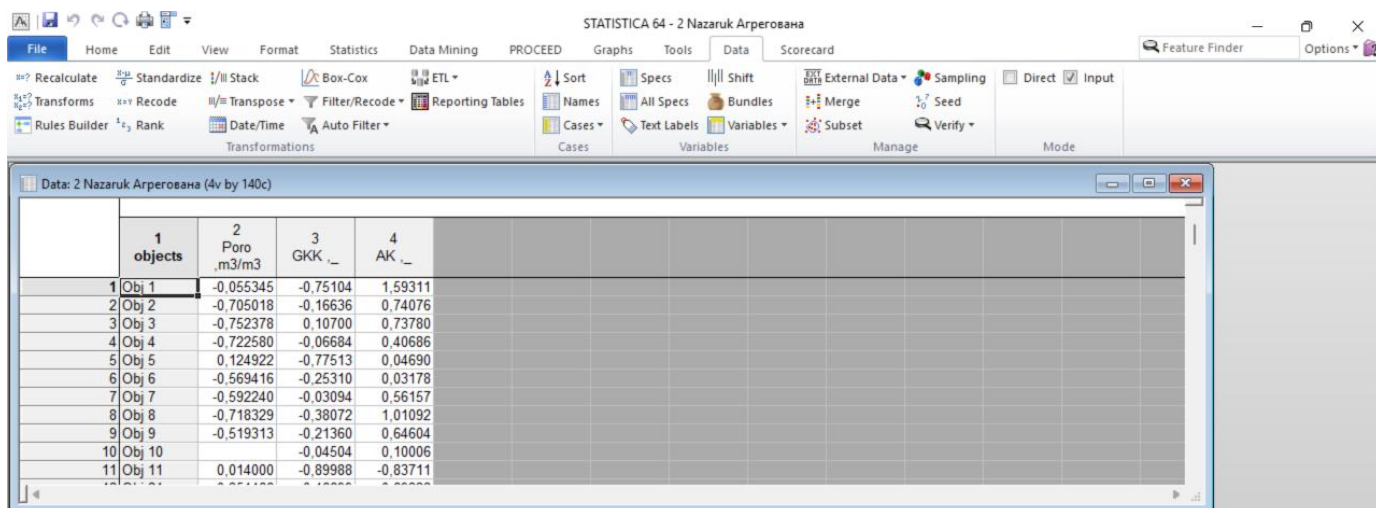
Такий підхід є найбільш математично коректним, так як дозволяє виконувати порівняння об'єктів у межах однакового набору параметрів. Якщо спільні ознаки достатньо добре описують фізичні властивості колекторів, саме в цьому просторі доцільно оцінювати ступінь спорідненості між пластами різних родовищ.

У межах першого типу моделювання до числа основних спільних ознак були включені:

- Пористість;
- Дані гамма-каротажу (GK);
- Інтервальна швидкість за акустичним каротажем (AK);

Для забезпечення сумісності даних попередньо було виконано перерахунок значень гамма-каротажу для перуанських свердловин із одиниць API у мкР/год. Це дозволило привести всі геофізичні параметри до єдиної системи вимірювання та уникнути спотворення результатів кластеризації.

Після формування простору спільних ознак дані по родовищах були об'єднані в єдину генеральну вибірку. Подальша підготовка та аналіз виконувалися в середовищі Statistica (Див.Рис.4.7).



	1 objects	2 Poro m3/m3	3 GK API	4 AK m/s
1	Obj 1	-0,055345	-0,75104	1,59311
2	Obj 2	-0,705018	-0,16636	0,74076
3	Obj 3	-0,752378	0,10700	0,73780
4	Obj 4	-0,722580	-0,06684	0,40686
5	Obj 5	0,124922	-0,77513	0,04690
6	Obj 6	-0,569416	-0,25310	0,03178
7	Obj 7	-0,592240	-0,03094	0,56157
8	Obj 8	-0,718329	-0,38072	1,01092
9	Obj 9	-0,519313	-0,21360	0,64604
10	Obj 10		-0,04504	0,10006
11	Obj 11	0,014000	-0,89988	-0,83711

Рис.4.7 - Агрегована таблиця геолого-геофізичних параметрів у середовищі Statistica

Оскільки початкові геофізичні параметри мали різну розмірність та різні діапазони значень, наступним етапом виконувалася стандартизація (Див.Рис.4.8) даних. Для цього використовувалася вбудована функція програми:

Data → Standardize.

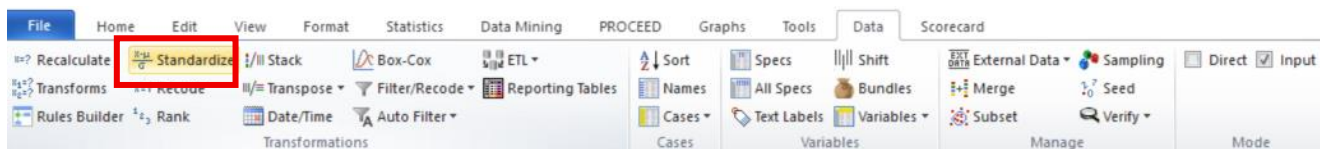


Рис.4.8 - Використання функції стандартизації даних

У результаті стандартизації всі параметри були переведені до z-оцінок, де середнє значення дорівнювало 0, а дисперсія — 1. Такий підхід дозволив усунути вплив різних масштабів параметрів та забезпечити коректне обчислення відстаней між об'єктами.

Особливістю вихідних матеріалів було те, що кількість каротажних записів значно перевищувала кількість самих пластів. У зв'язку з цим було виконано агрегування даних за допомогою функції Breakdown & One-way ANOVA, де в якості функції агрегації використовувалося середнє значення (Means). Це дозволило перейти від великої кількості окремих каротажних вимірювань до узагальнених характеристик кожного пласта.

Після агрегування було виконано механічне видалення пластів, які не містили достатньої кількості валідних геофізичних параметрів. У результаті була сформована фінальна агрегована вибірка, придатна для кластерного аналізу.

Безпосередньо кластеризація виконувалася у вигляді дерева кластеризації (Tree Clustering) у двох модифікаціях. У першому випадку в якості міри відстані використовувалася евклідова відстань (Euclidean distance), у другому — манхеттенська відстань (City-block або Manhattan distance).

В обох випадках для об'єднання кластерів застосовувався метод Ворда (Ward's method), основною перевагою якого є мінімізація внутрішньокластерної дисперсії та формування компактних груп об'єктів із близькими геофізичними характеристиками.

Таким чином, у межах першого типу моделювання було сформовано базову модель непараметричної класифікації пластів у просторі спільних геофізичних ознак.

Другий тип моделювання був спрямований на збереження максимальної кількості геолого-геофізичної інформації та врахування унікальних ознак кожної вибірки. На відміну від першого підходу, де використовувався лише простір спільних параметрів, у даному випадку застосовувався метод попарного виключення пропусків, який дозволяє виконувати кластеризацію навіть за наявності неоднорідних наборів ознак у різних об'єктах.

Початково у середовищі Statistica була сформована загальна таблиця геолого-геофізичних даних (Див.Рис.4.9), яка містила всі доступні параметри по пластах Полтавського ПСГ та перуанських родовищах.

	region	well	objects	Productive	Stratigr	DEPT,m	KNG _г	Poro _г ,m3/m3	VShale _г ,%	NGK _г	GKK _г	DSK	AK _г	BZ
1	ukraine	60	Obj 1			2600		-0,505965199	0,72165679	-0,702608538	-0,250727728	3,41703402	3,04386447	
2	ukraine	60	Obj 1			2604,90663		-0,531928953	-0,264275961	-0,603083049	-0,659035588	3,58274872	0,522632323	
3	ukraine	60	Obj 1			2624,88466		-0,102922701	-0,880968303	-0,644642158	-0,937621165	3,1032474	3,86309913	
4	ukraine	60	Obj 1			2644,86356		-0,807366304	0,302575406	-0,483300506	-0,498410707	2,72526814	0,911021889	
5	ukraine	60	Obj 1			2650		-0,315382422	-0,187668338	-0,634897865	-0,6203815	2,87674613	3,28825005	
6	ukraine	60	Obj 1			2664,84298		-0,698016351	0,505585132	-0,546627252	-0,458013815	2,4613344	1,77460562	
7	ukraine	60	Obj 1			2675		2,57871761	-1,6927892	-0,600226772	-1,50484063	2,0302953	0,734094174	
8	ukraine	60	Obj 1			2684,82368		-0,574807425	-0,127840497	-0,625017081	-0,600471907	2,90210169	1,85596914	
9	ukraine	60	Obj 1			2700		-0,617521494	0,72165679	-0,57575301	-0,282983842	2,49854466	1,15582195	
10	ukraine	60	Obj 1			2704,80658		-0,317934385	-0,439892918	-0,663898754	-0,717888121	3,4008023	3,02402866	
11	ukraine	60	Obj 1			2724,80093		-0,662557112	0,72165679	-0,569739174	-0,106797997	3,12762075	2,18006228	
12	ukraine	60	Obj 1			2743,42254		-0,715096577	0,636016132	-0,679054177	-0,372546251	3,37669893	1,7415169	
13	ukraine	60	Obj 1			2744,77408		-0,465220341	0,00486457373	-0,619413154	-0,569674668	2,85073148	2,61416647	
14	ukraine	60	Obj 1			2750		-0,396883957	-0,831560569	-0,489687052	-0,909717679	1,86820888	2,28691719	
15	ukraine	60	Obj 1			2764,7504		-0,45000219	-0,10596699	-0,64505581	-0,596464425	2,9264541	0,443388745	
16	ukraine	60	Obj 1			2775		-0,59064898	-0,860719183	-0,665897969	-0,925845736	2,69665035	1,06066233	
17	ukraine	60	Obj 1			2784,7265		-0,368481094	-0,791956585	-0,630190824	-0,887985301	2,79917179	2,42348754	
18	ukraine	60	Obj 1			2800		-0,598604665	-0,139872538	-0,632858389	-0,604253444	2,94878466	1,70298643	
19	ukraine	60	Obj 1			2804,70384		-0,558370194	-0,283780077	-0,711041513	-0,66409835	3,91229153	1,75837411	
20	ukraine	60	Obj 1			2824,68232		-0,730798288	-0,335407502	-0,519321233	-0,679996309	2,2951184	0,813245352	
21	ukraine	60	Obj 1			2844,6609		3,1103601	-2,09729628	-0,430364425	-1,86679313	0,607537317	0,732776803	
22	ukraine	60	Obj 1			2850		3,70641660	3,10551767	0,444010521	1,77440422	0,57151054	0,916776760	

Рис.4.9 - Початкова таблиця геолого-геофізичних параметрів для другого типу кластерного аналізу

Як і в попередньому методі, перед виконанням кластеризації було проведено стандартизацію параметрів за допомогою функції Data → Standardize. Після цього виконувалося агрегування каротажних записів через функцію Breakdown & One-way ANOVA із використанням середнього значення (Means) як функції агрегації. У результаті були отримані узагальнені характеристики окремих пластів (Див.Рис.4.10).

	1 objects	2 Poro ,m3/m3	3 GKK ,_	4 AK ,_
1	Obj 1	-0,055345	-0,75104	1,59311
2	Obj 2	-0,705018	-0,16636	0,74076
3	Obj 3	-0,752378	0,10700	0,73780
4	Obj 4	-0,722580	-0,06684	0,40686
5	Obj 5	0,124922	-0,77513	0,04690
6	Obj 6	-0,569416	-0,25310	0,03178
7	Obj 7	-0,592240	-0,03094	0,56157
8	Obj 8	-0,718329	-0,38072	1,01092
9	Obj 9	-0,519313	-0,21360	0,64604
10	Obj 10		-0,04504	0,10006
11	Obj 11	0,014000	-0,89988	-0,83711
12	Obj 21	0,351166	-0,16699	0,69332
13	Obj 22		0,08256	-0,44619
14	Obj 23	1,656938	-0,08685	0,06300
15	Obj 24	-0,204767	0,31192	0,34001
16	Obj 25	-0,762494	0,43407	0,41657
17	Obj 26	-0,155444	-0,20302	-0,40526
18	Obj 27	-0,539183	0,59477	-0,02891
19	Obj 28	-0,313165	0,04656	0,83604
20	Obj 29		0,64569	0,38442
21	Obj 30	0,421485	0,05606	0,23242
22	Obj 31	-0,434864	0,69560	0,24347
23	Obj 32	0,121935	-0,05074	-0,11612

Рис.4.10 - Агрегована вибірка геолого-геофізичних параметрів після стандартизації та узагальнення даних

Після агрегування було виконано механічне вилучення об'єктів, які не містили достатньої кількості валідних геофізичних параметрів.

Основна особливість даного підходу полягала у виборі міри відстані між об'єктами. Використання класичної евклідової відстані при різній кількості ознак могло призводити до спотворення результатів кластеризації. Для уникнення цього використовувалася матриця відстаней на основі показника: 1 – Pearson r.

Такий підхід дозволив оцінювати ступінь подібності профілів геофізичних характеристик між пластами незалежно від кількості спільних ознак, використовуючи лише наявні пари даних.

Кластеризація виконувалася у вигляді дерева кластеризації (Tree Clustering), а для об'єднання кластерів використовувався метод Ворда (Ward's method).

Таким чином, другий тип моделювання дозволив максимально зберегти вихідну геолого-геофізичну інформацію та врахувати унікальні характеристики окремих пластів.

Третім типом моделювання, використаним у даному дослідженні, стали мережі Кохонена (Kohonen SOM), які належать до методів навчання без учителя.

На відміну від класичних багат шарових персептронів (MLP), що потребують наявності правильних відповідей для навчання, мережі Кохонена призначені для кластеризації та пошуку подібності між об'єктами.

Практична реалізація моделі виконувалася у модулі SANN програмного середовища Statistica. Основною особливістю даного модуля є те, що нейронні мережі не допускають наявності пропущених значень у вхідній матриці даних. Тому перед початком моделювання було виконано додаткову підготовку даних.

Як і в першому типі моделювання, робота проводилася у просторі спільних геофізичних ознак, до яких були включені:

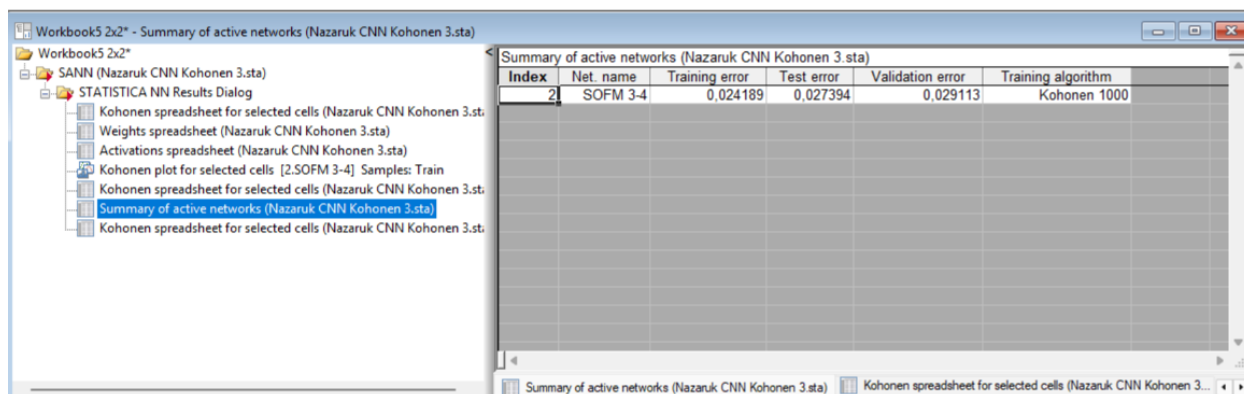
- Пористість;
- Дані гамма-каротажу (ГК);
- Дані акустичного каротажу (АК).

Наступним етапом виконувалося усунення пропущених значень методом імпутації. У межах окремого пласта відсутні значення замінювалися середнім значенням ознаки для даного пласта, а якщо ознака була відсутня повністю — використовувалося середнє значення для всієї свердловини. Це дозволило сформувати повну матрицю вхідних даних, придатну для роботи нейронної мережі (Див Рис.4.11).

	Var1	region	well	objects	Poro, m3/m3	GKK, _	AK, _
1	1	ukraine		60 Obj 1	0,0201999992	8,58300018	377
2	2	ukraine		60 Obj 1	0,0187314022	7,31716442	260,422363
3	3	ukraine		60 Obj 1	0,0429974347	6,4534936	414,880066
4	4	ukraine		60 Obj 1	0,0031517418	7,81513357	278,380859
5	5	ukraine		60 Obj 1	0,0309800003	7,4369998	388,299988
6	6	ukraine		60 Obj 1	0,0093369419	7,94037199	318,311554
7	7	ukraine		60 Obj 1	0,194680005	4,69500017	270,200012
8	8	ukraine		60 Obj 1	0,016306052	7,49872351	322,073669
9	9	ukraine		60 Obj 1	0,013890001	8,4829998	289,700012
10	10	ukraine		60 Obj 1	0,0308356527	7,13470984	376,082825
11	11	ukraine		60 Obj 1	0,0113426354	9,02921104	337,059204
12	12	ukraine		60 Obj 1	0,008370827	8,20533848	316,781586
13	13	ukraine		60 Obj 1	0,022504665	7,58420108	357,13147

Рис.4.11 - Підготовлена вхідна матриця геолого-геофізичних даних для навчання мережі Кохонена

Після підготовки даних виконувалося навчання мережі Кохонена (Див.Рис.4.12). Принцип роботи даного алгоритму полягає у проєктуванні багатовимірного простору геофізичних ознак на двовимірну топологічну карту нейронів. У результаті об'єкти з подібними геофізичними характеристиками потрапляють до одних і тих самих або сусідніх вузлів карти.



Index	Net. name	Training error	Test error	Validation error	Training algorithm
2	SOFM 3-4	0.024189	0.027394	0.029113	Kohonen 1000

Рис.4.12 - Результати навчання мережі Кохонена у середовищі Statistica

Оскільки наперед не було відомо оптимальної кількості кластерів, у роботі використовувався евристичний підхід із поступовим збільшенням розмірності карти Кохонена. Для цього послідовно використовувалися такі конфігурації мережі:

- $2 \times 2 = 4$ кластери;
- $2 \times 3 = 6$ кластерів;
- $3 \times 3 = 9$ кластерів;
- $3 \times 4 = 12$ кластерів;
- $4 \times 4 = 16$ кластерів;
- $4 \times 5 = 20$ кластерів;
- $5 \times 5 = 25$ кластерів.

Подальше збільшення кількості кластерів не виконувалося, оскільки починаючи приблизно з рівня 20 кластерів окремі групи вже були представлені лише 1–2 об'єктами, що ускладнювало геологічну інтерпретацію результатів.

Для оцінки розподілу об'єктів між нейронами використовувалися частотні карти та таблиці розподілу об'єктів по вузлах мережі (Див.Рис.4.13).

4.3 Аналіз результатів непараметричної класифікації та верифікація моделі на аналогах Перу

Отримані результати моделювання дозволяють оцінити ступінь подібності між пластами Полтавського ПСГ та Перуанських родовищ на основі їх геофізичних характеристик. Використання декількох підходів до непараметричної класифікації дало можливість порівняти поведінку об'єктів у різних моделях та простежити, наскільки стабільно окремі пласти формують спільні групи.

Це, у свою чергу, дозволяє більш обґрунтовано оцінити можливість використання окремих перуанських родовищ як аналогів при створенні цифрового близнюка Полтавського ПСГ.

Перший варіант кластерного аналізу виконувався у просторі спільних геофізичних ознак, тобто використовувалися лише ті параметри, які були наявні як для Полтавського ПСГ, так і для перуанських родовищ. До аналізу увійшло 110 об'єктів — окремих пластів (Див.Рис.4.14). Кластеризація проводилася методом Ворда, а для того щоб перевірити, наскільки результат залежить від способу розрахунку відстані між об'єктами, аналіз був виконаний у двох варіантах: з використанням евклідової та манхеттенської відстаней.

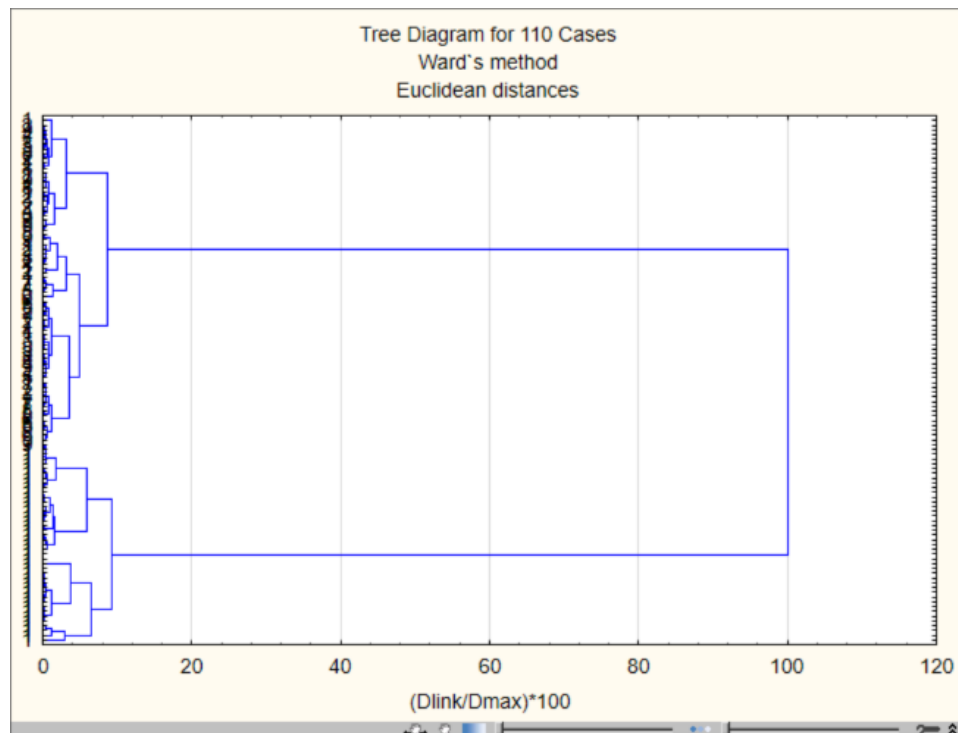


Рис.4.14 Дендрограма кластеризації пластів методом Ворда з використанням евклідової відстані

На дендрограмі для евклідової відстані видно, що більшість пластів об'єднуються між собою на відносно невеликих значеннях відстані. Це означає, що за основними геофізичними параметрами значна частина об'єктів має близькі характеристики. При цьому ближче до завершення кластеризації спостерігається різке об'єднання двох великих груп, що свідчить про помітну різницю між окремими частинами вибірки. Це добре видно і на графіку зміни відстаней об'єднання кластерів (Рис.4.15).

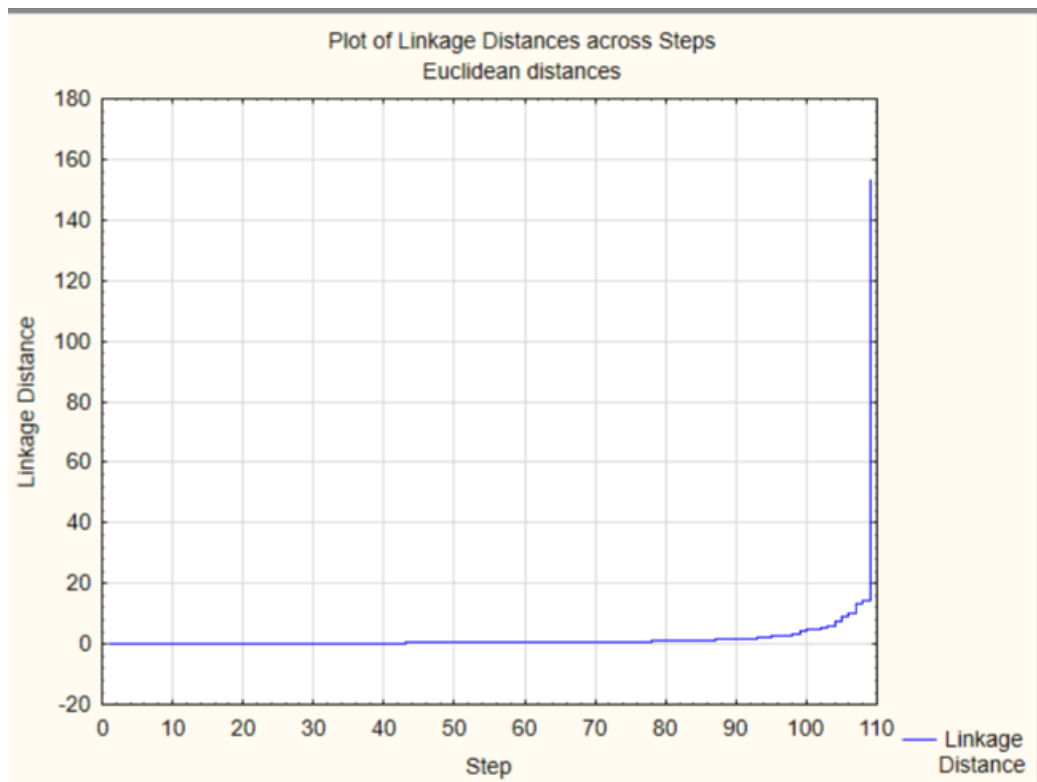


Рис.4.15 - Графік зміни відстаней об'єднання кластерів для евклідової метрики

На більшості етапів відстань між кластерами збільшується поступово, але наприкінці графіка спостерігається різкий стрибок. Це означає, що до певного моменту об'єднувалися дійсно схожі між собою пласти, а вже на фінальному етапі система змушена була поєднати між собою значно менш подібні групи об'єктів.

Другий варіант аналізу був виконаний із використанням манхеттенської відстані (Див.Рис.4.16).

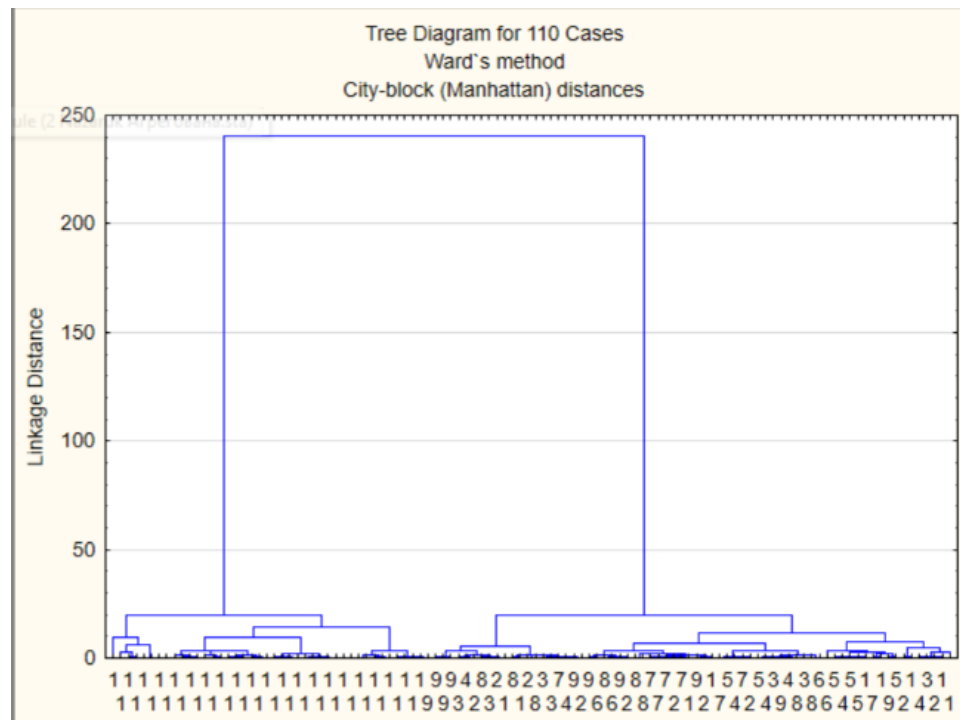


Рис.4.16 - Дендрограма кластеризації пластів методом Ворда з використанням манхеттенської відстані

Загальна структура дендрограми вийшла схожою, однак сам розподіл кластерів виглядає більш деталізованим. Частина пластів формує компактні групи, а частина відділяється на більших відстанях (Див.Рис.4.17). На мою думку, це пов'язано з тим, що манхеттенська відстань краще реагує на накопичення невеликих відмінностей між параметрами.

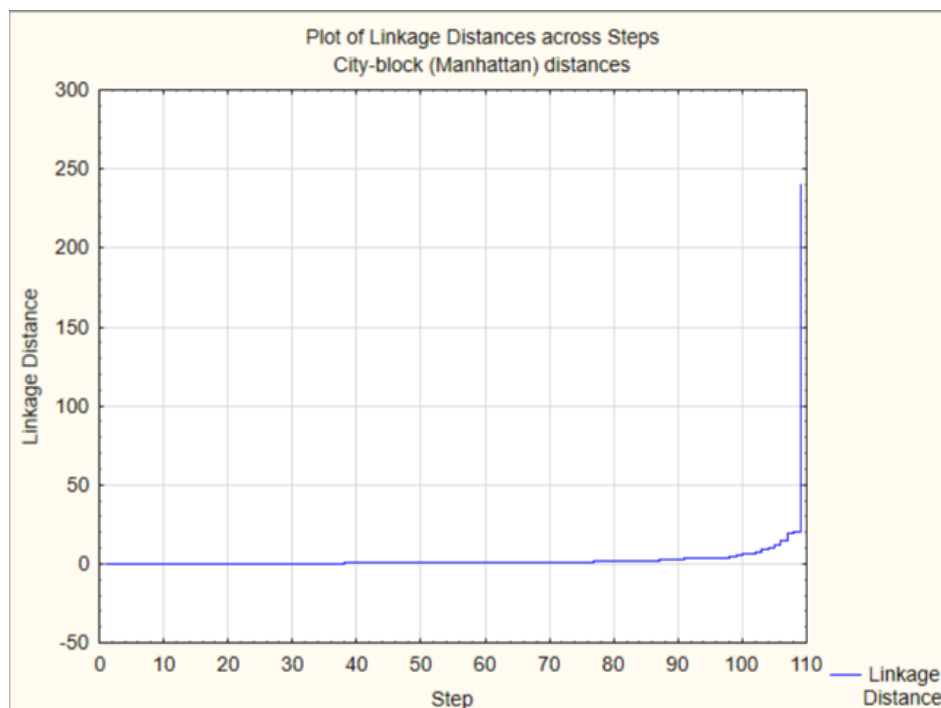


Рис.4.17 - Графік зміни відстаней об'єднання кластерів для манхеттенської метрики

Графік зміни відстаней для манхеттенської метрики також показує різке збільшення відстані на останніх етапах кластеризації. Це ще раз підтверджує, що всередині вибірки існують окремі групи пластів із близькими геофізичними характеристиками.

Обидва варіанти кластеризації дали подібну загальну картину, хоча манхеттенська відстань дозволила більш детально побачити внутрішню структуру вибірки. Це свідчить про те, що використані параметри — пористість, гамма-каротаж та акустичний каротаж — дійсно є достатньо інформативними для первинної оцінки подібності між пластами Полтавського ПСГ та перуанських родовищ.

Другий тип кластерного аналізу виконувався вже не лише у просторі спільних ознак, а з урахуванням максимально можливої кількості геолого-геофізичних параметрів. Основна ідея цього підходу полягала в тому, щоб не втрачати унікальні ознаки окремих пластів та свердловин, як це частково відбувалося у першому методі.

Для цього в якості міри подібності використовувався показник 1-Pearson r . На відміну від евклідової чи манхеттенської відстані, цей підхід оцінює не абсолютну різницю між параметрами, а схожість форми геофізичного профілю об'єктів. Тобто важливим стає не саме значення параметра, а характер зміни ознак між пластами.

На дендрограмі видно, що структура кластерів стала більш складною та розгалуженою порівняно з першим методом (Див.Рис.4.18). Якщо у випадку евклідової та манхеттенської відстаней вибірка досить швидко розділялася на дві великі групи, то тут формування кластерів відбувається більш поступово.

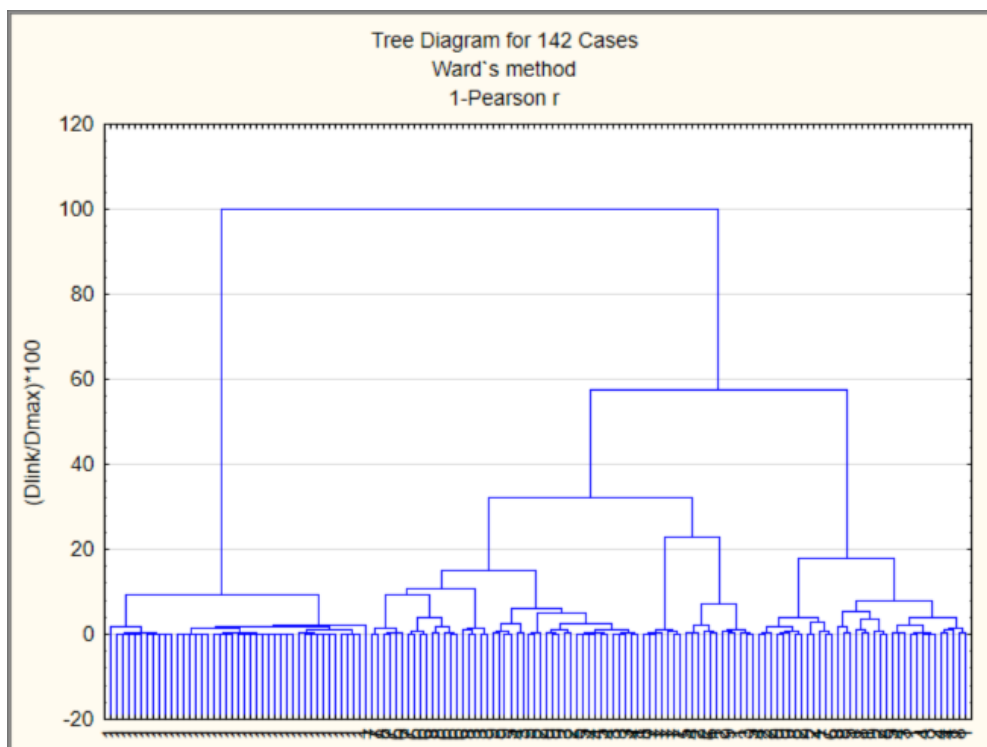


Рис.4.18 - Дендрограма кластеризації пластів методом Ворда з використанням міри подібності 1-Pearson r

Це може свідчити про те, що частина пластів має подібний характер зміни геофізичних параметрів навіть при різних абсолютних значеннях. Іншими словами, окремі об'єкти можуть бути схожими не за конкретними величинами пористості чи ГК, а саме за співвідношенням параметрів між собою.

Також помітно, що в середині вибірки формується декілька окремих підгруп, які об'єднуються між собою лише на більших рівнях відстані. На мою думку, це

говорить про більш складну внутрішню структуру даних та наявність декількох типів геолого-геофізичних профілів серед досліджуваних пластів.

Графік зміни відстаней об'єднання кластерів також виглядає інакше, ніж у першому методі. На початкових етапах зміна відстаней практично не помітна, а різке зростання починається лише ближче до завершення кластеризації (Див.Рис4.19). Це означає, що значна частина об'єктів має відносно близький характер геофізичних профілів і об'єднується між собою без різких стрибків відстані.

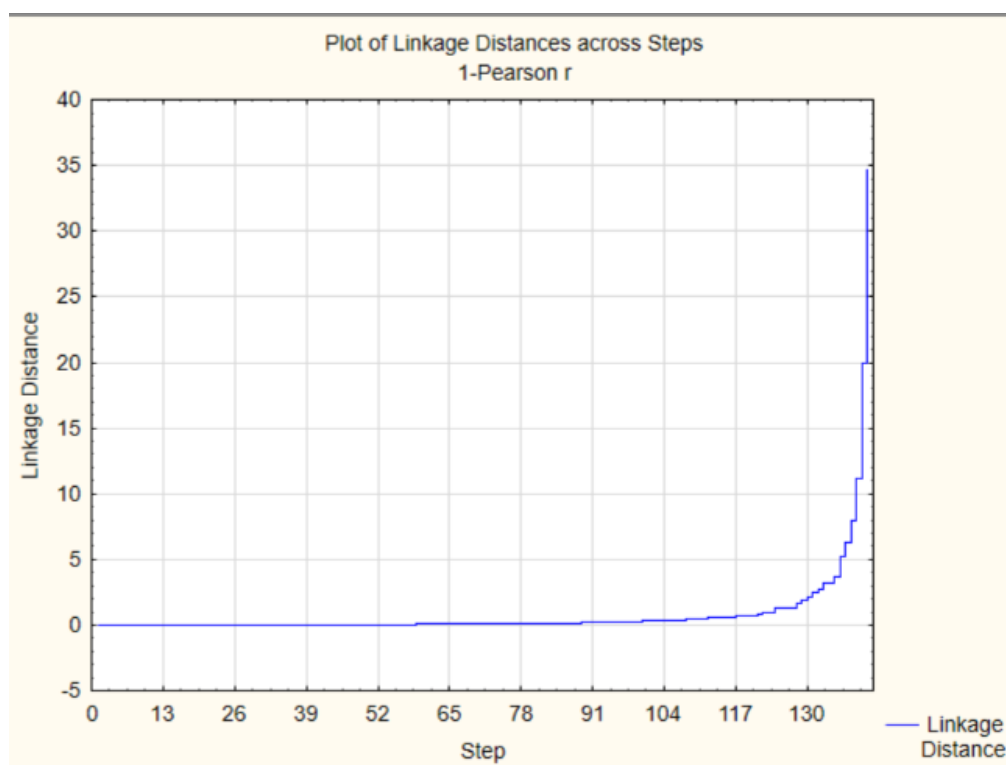


Рис.4.19 - Графік зміни відстаней об'єднання кластерів для міри подібності 1-Pearson r

Лише на останніх етапах відбувається об'єднання вже суттєво різних груп пластів, через що графік демонструє різкий підйом. Така поведінка є доволі важливою, оскільки свідчить про те, що вибірка не є хаотичною, а містить внутрішню структурованість.

Порівняно з першим методом, кластеризація на основі 1-Pearson r дала більш гнучкий результат. Метод дозволив врахувати значно більшу кількість вихідної інформації та уникнути жорсткого відсікання ознак через пропуски в даних. У результаті структура кластерів вийшла більш деталізованою, а сам розподіл пластів — менш різким і більш природним.

Третій метод аналізу був виконаний із використанням мереж Кохонена, тобто самоорганізованих карт, які дозволяють не просто розділити об'єкти на групи, а показати, які пласти за своїми геофізичними характеристиками «притягуються» до одних і тих самих нейронів. У цьому випадку основна увага була зосереджена не на абсолютному поділі вибірки, а на тому, чи потрапляють українські та перуанські об'єкти в однакові або близькі нейронні комірки. Саме це є важливим для перевірки гіпотези подібності.

На загальному графіку активації для всіх груп видно, що більшість об'єктів належить до української вибірки. Це очікувано, оскільки кількість записів по Полтавському ПСГ у вхідному наборі більша (Див.Рис.4.20).

Перуанські об'єкти представлені меншою кількістю точок і на графіку виділяються окремою червоною лінією. Важливо, що вони не відокремлюються повністю від українських даних, а потрапляють у той самий простір активацій. Це означає, що нейромережа не сприйняла перуанські об'єкти як повністю чужорідні, а знайшла для них місце в загальній структурі даних.

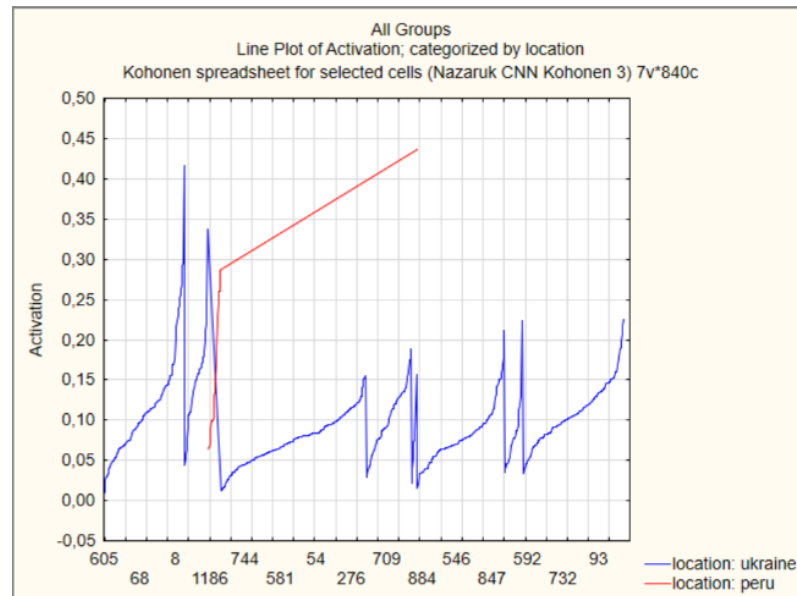


Рис.4.20 - Загальний графік активації об'єктів української та перуанської вибірок у мережі Кохонена

Під час аналізу окремих нейронних комірок видно, що більшість нейронів заповнена переважно українськими об'єктами. Наприклад, у вузлах з координатами (1,1), (1,2), (1,3), (2,1), (2,2), (3,1), (3,2) та (3,3) переважають саме об'єкти Полтавського ПСГ. Це свідчить про те, що всередині української вибірки існує власна внутрішня структура: різні пласти мають відмінні комбінації пористості, гамма-каротажу та акустичного каротажу, тому вони розподіляються між різними нейронами.

Особливо показовим є вузол (2,3). На цьому графіку, крім українських об'єктів, присутній також перуанський об'єкт, який потрапив до тієї самої нейронної комірки. Це дуже важливий результат, оскільки саме такі випадки показують потенційну геофізичну подібність між окремими пластами Полтавського ПСГ та перуанськими аналогами(Див.Рис.4.21).

Тобто мережа Кохонена, працюючи без попередньо заданих «правильних» відповідей, самостійно віднесла частину перуанських даних до тієї ж групи, що й українські об'єкти.

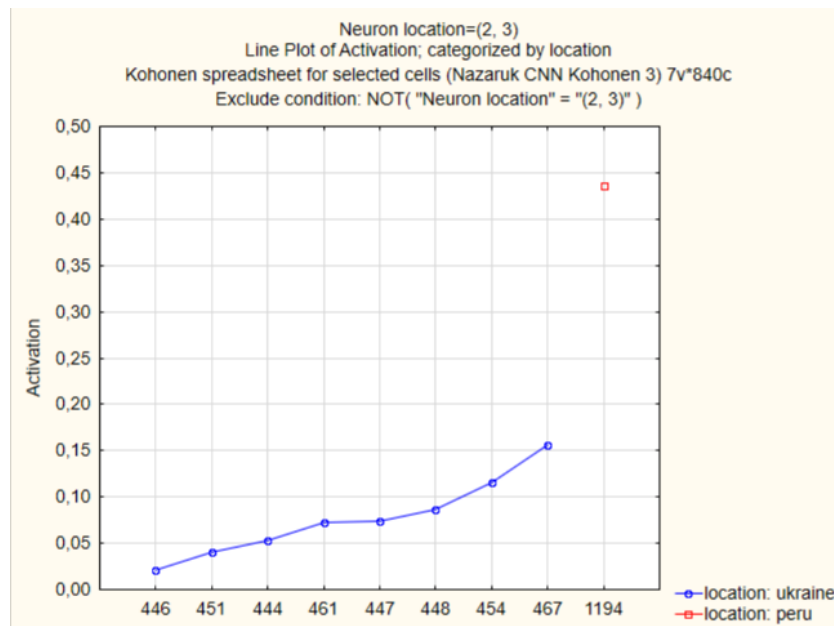


Рис.4.21 - Графік активації для нейрона (2,3), у якому спостерігається спільне розміщення українського та перуанського об'єктів

Це не означає, що всі перуанські родовища можна однозначно вважати аналогами Полтавського ПСГ. Навпаки, результати показують, що подібність має вибірковий характер. Більшість перуанських об'єктів не формує великої спільної групи з українськими пластами, але окремі об'єкти все ж потрапляють у спільні або близькі нейронні зони. Саме це і є найбільш реалістичним результатом для геологічних даних, оскільки повного збігу між різними басейнами очікувати не варто.

Якщо порівнювати цей результат із попередніми кластерними методами, мережа Кохонена дала більш гнучке представлення структури даних. Класична кластеризація формує жорстке дерево, де кожен об'єкт поступово об'єднується з іншими.

Натомість SOM дозволяє побачити, як об'єкти розподіляються по нейронній сітці, і чи є між ними локальні зони подібності. Саме тому цей метод виявився корисним для пошуку не загальної, а точкової подібності між українськими та перуанськими пластами.

Отже, результати всіх трьох підходів до непараметричної класифікації показали, що методи кластерного аналізу та мережі Кохонена є доцільними для оцінки подібності геолого-геофізичних об'єктів. Вони дозволили виявити внутрішню структуру вибірок та оцінити ступінь близькості між пластами Полтавського ПСГ і перуанських родовищ у просторі геофізичних ознак.

Разом з тим, отримані результати не підтвердили наявності стійкої та вираженої подібності між більшістю пластів української та перуанської вибірок. У більшості випадків об'єкти формували окремі групи, а випадки їх спільного об'єднання були поодинокими. Аналогічна ситуація спостерігалася і при використанні мереж Кохонена, де лише окремі перуанські об'єкти потрапляли до спільного нейронного простору з пластами Полтавського ПСГ.

На мою думку, це свідчить про те, що гіпотеза повної геофізичної подібності між досліджуваними об'єктами загалом не підтвердилася. Водночас окремі пласти все ж демонструють близькі геофізичні характеристики, що дозволяє розглядати їх як локальні аналоги при побудові цифрових моделей.

ВИСНОВКИ

У процесі виконання магістерської роботи було проаналізовано геологічну будову та геолого-геофізичні особливості Полтавського підземного сховища газу.

Встановлено, що досліджуваний об'єкт характеризується складною багатоблоковою будовою, неоднорідністю пластів-колекторів та значною варіативністю петрофізичних параметрів, що ускладнює процес створення цифрового близнюка.

У роботі було розглянуто сучасні підходи до створення цифрових близнюків у геофізичних дослідженнях, а також проаналізовано можливості використання методів кластер-аналізу та самоорганізовуваних мереж Кохонена для роботи зі складними багатовимірними геоданими.

Показано, що ефективність таких моделей значною мірою залежить від якості підготовки, структурування та інтеграції геолого-геофізичної інформації.

У ході дослідження було сформовано багатопараметричну вибірку геолого-геофізичних даних для Полтавського ПСГ та ряду нафтогазових родовищ Перу. Для аналізу використовувалися дані гамма-каротажу, електричного опору, густинного й акустичного каротажу, а також основні петрофізичні характеристики пластів-колекторів. Після підготовки та уніфікації даних було проведено кластеризацію та аналіз подібності між досліджуваними об'єктами.

Отримані результати показали, що між Полтавським ПСГ та досліджуваними перуанськими родовищами.

Не було виявлено, принаймі на рівні формальних математичних закономірностей, що демонстрували б достатній рівень комплексної подібності за сукупністю геолого-геофізичних параметрів. Можна стверджувати, що регіональні відмінності пов'язані з літологічною будовою, структурними особливостями колекторів та умовами формування резервуарів є надзвичайно значними, що не

дозволяє підтвердити гіпотезу подібності геофізичних характеристик і таким чином мати варіант для виокремлення «цифрового близнюка».

Разом із тим виконана робота показала ефективність використання кластер-аналізу та самоорганізовуваних мереж Кохонена для аналізу складних багатовимірних геоданих і підтвердила перспективність їх застосування при створенні цифрових моделей та цифрових близнюків підземних сховищ газу.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

Вижва, С. А., Безродна, І. М., Козіонова, О. О. (2012). *Аналіз ємнісних властивостей карбонатних порід нижнього карбону Руденківсько-Пролетарського НГР.*

https://www.researchgate.net/publication/301433353_Analiz_emnysnih_vlastiv_oste_j_karbonatnih_porid_karbonu_Rudenkivsko-Proletarskoi_NGO_za_rezultatami_GDS_ta_petrofiziki

Гілецький, Й. Р., Сливка, Р. Р., & Богович, М. М. (2008). *Географія: довідник*. Веста; Ранок.

https://pidru4niki.com/12991010/geografiya/relyef_geolopchna_budova_korish_i_kopalini

Закревский, К. Е., Попов, В. Л. (2021). *История развития трехмерного геологического моделирования как метода изучения залежей нефти и газа*. Bulletin of the Tomsk Polytechnic University. Geo Assets Engineering, 332(5), 89–100.

<https://www.researchgate.net/publication/352008203>

І. Осадчий та ін. (2023). *Нетрадиційні газові ресурси у східному нафтовидовувальному регіоні України*. Вісник Київського національного університету імені Тараса Шевченка. Геологія, 2(101), 100–107.

<https://geology.bulletin.knu.ua/en/article/view/1763>

Коболев, А. В., Федутенко, А. М., Шумко, Р. Й., Локтев, С. А. (2018). *Сучасний стан та перспективи розвитку підземних сховищ газу України*. Нафтогазова галузь України, (3), 3–10.

Роздорожний, В. І. (2015). *Еволюція нафтогазової інфраструктури України: історико-технічний нарис*. Техніка.

Укртрансгаз. (2025). *Геологічна характеристика Полтавського підземного сховища газу: фондові матеріали (внутрішній документ)*. Київ.

Харківський національний університет імені В. Н. Каразіна. (2024).

Геофізика: навчально-методичний комплекс.

https://geology.univer.kharkov.ua/wp-content/uploads/2024/01/NMK_Heofizyka.pdf

Abdo, E., Al-Kuwari, K., Al-Thani, A. A., Al-Kulaifi, A. J., & Al-Hajri, M. S. (2023). *Case study of the use of a digital twin for leak detection and integrity monitoring*. SPE Journal, 28(05), 2415–2433.

<https://onepetro.org/SJ/article-abstract/28/05/2415/520757/Case-Study-of-the-Use-of-a-Digital-Twin-for-Leak>

Abdo, E., Baronio, E., Mauro, S., Troise, M., & Salamina, L. (2023). *Case study of the use of a digital twin for leak detection and quantification in underground gas storage wells*. Politecnico di Torino.

<https://iris.polito.it/handle/11583/2979377>

Athibodhi, N., Mohanty, S., Gupta, P. J., Al-Mulla, A., Al-Jabri, M., & Al-Qubaisi, M. (2023). *Production optimization and enhancement through digital twin – A field case study*. Abu Dhabi International Petroleum Exhibition and Conference.

<https://onepetro.org/SPEADIP/proceedings-abstract/23ADIP/23ADIP/D031S108R007/534614>

Baçaõ, F., Lobo, V., & Painho, M. (2005). The self-organizing map, the Geo-SOM, and relevant variants for geosciences. *Computers & Geosciences*, 31(2), 155–163.

<https://www.sciencedirect.com/science/article/abs/pii/S0098300404001918>

Bamberg, A., Urbas, L., Bröcker, S., Bortz, M., & Kockmann, N. (2021). *The digital twin – Your ingenious companion for process engineering and smart production*. Chemical Engineering & Technology.

<https://www.researchgate.net/publication/349634634>

Bergen, K. J., Johnson, P. A., de Hoop, M. V., & Beroza, G. C. (2019). *Machine learning for data-driven discovery in solid Earth geoscience*. Science, 363.

<https://www.science.org/doi/10.1126/science.aau0323>

Bestagini, P., Lipari, V., & Tubaro, S. (2017). *A machine learning approach to facies classification using well logs*. SEG Technical Program Expanded Abstracts 2017, 2137–2142.

https://www.researchgate.net/publication/319171157_A_machine_learning_approach_to_facies_classification_using_well_logs

Breiman, L., Friedman, J., Olshen, R. A., & Stone, C. J. (1984). *Classification and Regression Trees*. Routledge.

<https://www.taylorfrancis.com/books/mono/10.1201/9781315139470/classification-regression-trees-leo-breiman-jerome-friedman-olshen-charles-stone>

Chernova, O. T. (2014). *The importance of underground storage on the world market and tendencies of underground gas storage development*. Exploration and Development of Oil and Gas Fields, (4), 18–25.

Corrales, D. C., Ledezma, A. I., & Corrales, J. C. (2015). *A conceptual framework for data quality in knowledge discovery tasks (FDQ-KDT): A proposal*.

https://www.researchgate.net/figure/Data-science-process-flowchart_fig3_283430974

Cottrell, M., Verleysen, M. (2006). *Advances in Self Organising Maps*.

<https://arxiv.org/abs/cs/0611058>

Dehghani, M., & Azadifar, M. (2016). *Incremental learning for large-scale churn prediction*. ResearchGate.

https://www.researchgate.net/publication/306136478_Incremental_Learning_for_Large_Scale_Churn_Prediction

Demchuk, Y., Shogenov, K., Shogenova, A., Merson, B., & Vincent, C. J. (2025). *Geological and petrophysical properties of underground gas storage facilities in Ukraine and their potential for hydrogen and CO₂ storage*. Sustainability, 17(6), Article 2400.

<https://www.mdpi.com/2071-1050/17/6/2400>

Du, H., Cao, J., Xue, Y., & Wang, X. (2015). Seismic facies analysis based on self-organizing map and empirical mode decomposition. *Journal of Applied Geophysics*, 112, 52–61.

<https://www.sciencedirect.com/science/article/abs/pii/S0926985114003267>

Emerick, A. A., & Reynolds, A. C. (2012). *History matching time-lapse seismic data using the ensemble Kalman filter with multiple data assimilations*. *Computational Geosciences*, 16(3), 639–659. <https://link.springer.com/article/10.1007/s10596-012-9275-5>

Evans, D. J., & Chadwick, R. A. (Eds.). (2009). *Underground gas storage: Worldwide experiences and future development in the UK and Europe*. Geological Society of London.

[https://www.researchgate.net/publication/274575879 Underground Gas Storage Worldwide Experiences and Future Development in the UK and Europe](https://www.researchgate.net/publication/274575879_Underground_Gas_Storage_Worldwide_Experiences_and_Future_Development_in_the_UK_and_Europe)

Gahlot, A. P., Gupta, A., Vishal, V., & Singh, L. (2024). *A digital twin for geological carbon storage with controlled injectivity*. *arXiv*.

<https://arxiv.org/abs/2403.19819>

García, S., Luengo, J., & Herrera, F. (2015). *Data preprocessing in data mining*. Springer.

<https://link.springer.com/book/10.1007/978-3-319-10247-4>

George, A. S. (2024). *The impact of IT/OT convergence on digital transformation in manufacturing*.

[https://www.researchgate.net/publication/379406907 The Impact of ITOT Convergence on Digital Transformation in Manufacturing](https://www.researchgate.net/publication/379406907_The_Impact_of_ITOT_Convergence_on_Digital_Transformation_in_Manufacturing)

Glaessgen, E., & Stargel, D. (2012). *The digital twin paradigm for future NASA and U.S. Air Force vehicles*.

<https://ntrs.nasa.gov/citations/20120008178>

Grieves, M. (2014). *Digital twin: Manufacturing excellence through virtual factory replication*.

https://www.researchgate.net/publication/275211047_Digital_Twin_Manufacturing_Excellence_through_Virtual_Factory_Replication

Guarido, M. (2019). *Machine learning strategies to perform facies classification*. GeoConvention 2019.

https://geoconvention.com/wp-content/uploads/abstracts/2019/GC2019_161_Machine_Learning_Strategies_to_Perform_Facies_Classification.pdf

Habib, M., Yao, G., Xie, C., Charles, S. P., Jakada, H., Danlami, M. S., Ahmed, H. A., & Ibrahim, A. O. (2017). *Optimizing oil and gas field management through a fractal reservoir study model*. Journal of Petroleum Exploration and Production Technology, 7(1), 43–53.

<https://link.springer.com/article/10.1007/s13202-016-0252-8>

Hall, M. (2016). *Facies classification using machine learning*. The Leading Edge, 35(10), 906–909.

<https://pubs.geoscienceworld.org/seg/tle/article/35/10/906/292308/Facies-classification-using-machine-learning>

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction (2nd ed.)*. Springer.

<https://link.springer.com/book/10.1007/978-0-387-84858-7>

Huang, F., Xiong, H., Yao, C., Catani, F., Zhou, C., & Huang, J. (2018). *Predicting landslide susceptibility considering different landslide types*. Journal of Rock Mechanics and Geotechnical Engineering.

https://www.researchgate.net/figure/Predict-the-susceptibility-and-landslide-type-separately_fig1_325524542

Iwaszczuk, N., et al. (2022). *Underground Gas Storage Facilities in Ukraine: Current State and Future Prospects*. Energies, 15(18), 6604.

<https://www.mdpi.com/1996-1073/15/18/6604>

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning: With Applications in R*. Springer.

<https://link.springer.com/book/10.1007/978-1-4614-6849-3>

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An Introduction to Statistical Learning: With Applications in R (2nd ed.)*. Springer.

<https://www.statlearning.com/>

Jones, D., Snider, C., Nassehi, A., Yon, J., & Hicks, B. (2020). *Characterising the digital twin: A systematic literature review*. CIRP Journal of Manufacturing Science and Technology, 29, 36–52.

<https://www.sciencedirect.com/science/article/pii/S1755581720300110>

Karpatne, A., Ebert-Uphoff, I., Ravela, S., Babaie, H. A., & Kumar, V. (2017). *Machine learning for the geosciences: Challenges and opportunities*. arXiv.

<https://arxiv.org/abs/1711.04708>

Kempler, S., & Mathews, T. (2017). *Earth science data analytics: Definitions, techniques and skills*. Data Science Journal, 16, 6.

<https://datascience.codata.org/articles/10.5334/dsj-2017-006>

Kohonen, T. (1982). *Self-organized formation of topologically correct feature maps*. Biological Cybernetics, 43(1), 59–69.

<https://link.springer.com/article/10.1007/BF00337288>

Kohonen, T. (2001). *Self-organizing maps* (3rd ed.). Springer.

<https://link.springer.com/book/10.1007/978-3-642-56927-2>

Kritzinger, W., Karner, M., Traar, G., Henjes, J., & Sihn, W. (2018). *Digital twin in manufacturing: A categorical literature review and classification*. IFAC-PapersOnLine, 51(11), 1016–1022.

<https://www.sciencedirect.com/science/article/pii/S2405896318316021>

Kuhn, M., & Johnson, K. (2013). *Applied Predictive Modeling*. Springer.

<https://link.springer.com/book/10.1007/978-1-4614-6849-3>

Li, X., Zhang, Y., Wang, Z., & Chen, G. (2020). *Digital twin-driven smart reservoir management: Framework and case study*. Computers & Geosciences, 145, 104576.

<https://www.sciencedirect.com/science/article/abs/pii/S0166361520305546>

Liu, W., Chen, Z., Hu, Y., & Xu, L. (2022). *A systematic machine learning method for reservoir identification and production prediction*. Petroleum Science, 20(1), 295–308.

<https://www.cup.edu.cn/petroleumscience/docs/2023-03/dc4e0ffe28524cd3b2b1424026c751bd.pdf>

Ma, Y. Z. (2011). *Uncertainty analysis in reservoir characterization and management: How much should we know about what we don't know?* American Association of Petroleum Geologists.

<https://store-assets.aapg.org/documents/previews/1097M96/CHAPTER01.pdf>

Meza, E. B. M., Souza, D. G. B. d., Copetti, A., Sobral, A. P. B., Silva, G. V., Tammela, I., & Cardoso, R. (2024). *Tools, technologies and frameworks for digital twins in the oil and gas industry: An in-depth analysis*. Sensors, 24(19), 6457.

<https://www.mdpi.com/1424-8220/24/19/6457>

Molnar, C. (2022). *Interpretable machine learning: A guide for making black box models explainable* (2nd ed.).

<https://christophm.github.io/interpretable-ml-book/>

Narayan, N. S., Dwivedi, R., & Maurya, S. P. (2023). *Machine learning for lithofacies classification using well log and seismic data: A review*. Earth-Science Reviews, 245, 104543.

Peeters, L., Dassargues, A. (2006). *Comparison of Kohonen's Self-Organizing Map Algorithm and Standard Clustering Methods*.

<https://orbi.uliege.be/bitstream/2268/3388/1/publi144-2007.pdf>

Provost, F., & Fawcett, T. (2013). *Data science for business: What you need to know about data mining and data-analytic thinking*. O'Reilly Media.

<https://www.oreilly.com/library/view/data-science-for/9781449374273>

Rahimi, M., & Riahi, M. A. (2022). *Reservoir facies classification based on random forest and geostatistics methods in an offshore oilfield*. Journal of Applied Geophysics, 201, 104640.

<https://www.sciencedirect.com/science/article/abs/pii/S0926985122001112>

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat. (2019). *Deep learning and process understanding for data-driven Earth system science*. Nature, 566(7743), 195–204.

<https://www.nature.com/articles/s41586-019-0912-1>

Sadri, M., Shariatipour, S. M., Hunt, A., & Ahmadinia, M. (2019). *Effect of systematic and random flow measurement errors on history matching: A case study on oil and wet gas reservoirs*.

<https://www.researchgate.net/publication/302918629>

Sagir, A. M., & Sathasivam, S. (2016). The self-organizing map as a tool for cluster analysis. Journal of Applied Sciences, 16(3), 115–125.

<https://www.researchgate.net/publication/357869973> The Self Organizing Map as a Tool for Cluster Analysis

Sokolova, M., & Lapalme, G. (2009). *A systematic analysis of performance measures for classification tasks*. Information Processing & Management, 45(4), 427–437.

<https://www.sciencedirect.com/science/article/abs/pii/S0306457309000259>

Smith, T., Treitel, S. (2011). *Introduction to Self-Organizing Maps in Multi-Attribute Seismic Data*.

<https://www.geoinsights.com/introduction-to-self-organizing-maps-in-multi-attribute-seismic-data>

Tao, F., Cheng, J., & Qi, Q. (2018). *Digital twin-driven product design, manufacturing and service with big data*. International Journal of Advanced Manufacturing Technology.

<https://www.researchgate.net/publication/315328118> Digital twin-driven product design manufacturing and service with big data

Tao, F., Zhang, H., Liu, A., & Nee, A. Y. C. (2019). *Digital twin in industry: State-of-the-art*. *IEEE Transactions on Industrial Informatics*, 15(4), 2405–2415.

<https://ieeexplore.ieee.org/document/8477101>

Tyasi, T. L., Eydurán, E., & Celik, S. (2021). *Comparison of tree-based regression tree methods for predicting live body weight from morphological traits in Hy-line silver brown commercial layer and indigenous Potchefstroom Koekoek breeds raised in South Africa*. *Tropical Animal Health and Production*, 53(7).

https://www.researchgate.net/figure/Descriptive-statistics-of-body-weight-and-morphological-traits_tbl2_344617986

Veerman, M. A., Pincus, R., Stoffer, R., van Leeuwen, C. M., Podareanu, D., & van Heerwaarden, C. C. (2021). *Predicting atmospheric optical properties for radiative transfer computations using neural networks*. *Philosophical Transactions of the Royal Society A*.

<https://royalsocietypublishing.org/rsta/article/379/2194/20200095/41186/Predicting-atmospheric-optical-properties-for>

Yang, Y., Peng, J., Xu, T., Wang, Y., & Zeng, Y. (2022). *Characterization, classification, and evaluation of reservoir pore structure features of lacustrine fine-grained sedimentary rocks*. *Frontiers in Earth Science*, 10, 878089.

<https://www.frontiersin.org/journals/earth-science/articles/10.3389/feart.2022.878089/full>

Yilmaz, O. (2024). *Detection and prediction of hydrogen leakage in salt caverns using digital twin technology*. *Proceedings of the 43rd International Conference on Ocean, Offshore and Arctic Engineering*.

<https://asmedigitalcollection.asme.org/OMAE/proceedings-abstract/OMAE2024/87868/V008T11A063/1202831>

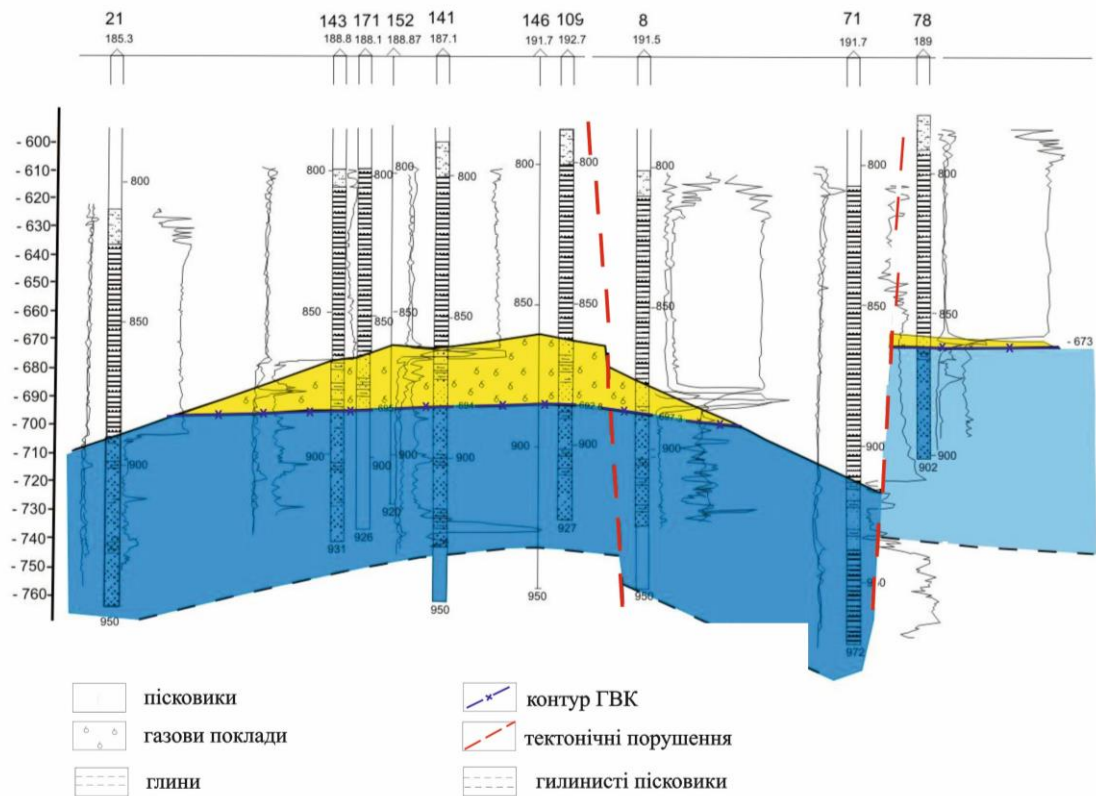
Zhang, T., Li, Y., Cai, J., Meng, Q., Sun, S., & Li, C. (2020). *A digital twin for unconventional reservoirs: A multiscale modeling and algorithm to investigate complex mechanisms*. *Geofluids*, Article 8876153.

<https://onlinelibrary.wiley.com/doi/10.1155/2020/8876153>

ДОДАТКИ

Додаток А

Поперечний геолого-літологічний профіль



Додаток Б

Структурна карта покрівлі горизонту J_{2b}

