

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет _____ Комп'ютерних наук
(повна назва)
Кафедра _____ Штучного інтелекту
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

рівень вищої освіти _____ перший (бакалаврський)

Згорткова нейронна мережа для оцінювання показника Херста
фрактального броунівського руху
(тема)

Виконав:
здобувач _____ четвертого _____ року навчання,
групи _____ ІТШ-22-1

_____ Софія ЛАВРИНЕНКО
(власне ім'я, прізвище)

Спеціальність 122 Комп'ютерні науки

(код і повна назва спеціальності)

Тип програми _____ освітньо-професійна

Освітня програма _____ Штучний інтелект

(повна назва освітньої програми)

Керівник _____ проф. Лариса ЧАЛА
(посада, власне ім'я, прізвище)

Допускається до захисту

Завідувач кафедри ШІ

(підпис)

_____ Лариса ЧАЛА
(власне ім'я, прізвище)

2026 р.

Харківський національний університет радіоелектроніки

Факультет _____ Комп'ютерних наук _____

Кафедра _____ Штучного інтелекту _____

Рівень вищої освіти _____ перший (бакалаврський) _____

Спеціальність _____ 122 Комп'ютерні науки _____
(код і повна назва)

Тип програми _____ освітньо-професійна _____

Освітня програма _____ Штучний інтелект _____
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____

(підпис)

« _____ » _____ 20 ____ р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

здобувачеві _____ Лавриненко Софії Романівні _____
(прізвище, ім'я, по батькові)

1. Тема роботи Згортова нейронна мережа для оцінювання показника Херста
фрактального броунівського руху

затверджена наказом університету від 5 червня 2026 р. № 448Ст

2. Термін подання здобувачем роботи до екзаменаційної комісії 9 липня 2026 р.

3. Вихідні дані до роботи Дані інтернет-джерел, середовище розробки Google Colab,
науково-технічні публікації, мовні моделі, датасети аномальної дифузії AnDi, Python-
бібліотеки.

4. Перелік питань, що потрібно опрацювати в роботі _____

1) Аналіз процесу оцінювання показника Херста для часових рядів фрактального
броунівського руху

2) Методи оцінювання показника Херста для фрактального гауссового шуму

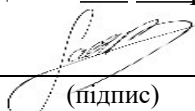
3) Реалізація оцінювача показника Херста на основі CNN та експериментальні дослідження

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Строк / термін виконання етапів роботи	Примітка
1	Отримання завдання на кваліфікаційну роботу	08.06.2026	виконано
2	Аналіз предметної галузі та постановка задачі	08.06.2026 – 09.06.2026	виконано
3	Огляд методів оцінки показника Херста часових рядів фрактального броунівського руху	09.06.2026 – 10.06.2026	виконано
4	Проектування архітектури системи синтезу навчальної вибірки	11.06.2026 – 13.06.2026	виконано
5	Реалізація базового оцінювача показника Херста на основі згорткової нейромережі	13.06.2026 – 15.06.2026	виконано
6	Розробка плану досліджень	15.06.2026 – 16.06.2026	виконано
7	Налаштування експериментального середовища	17.06.2026 – 18.06.2026	виконано
8	Проведення експериментів та візуалізації	18.06.2026 – 21.06.2026	виконано
9	Оцінювання результатів експериментів	21.06.2026 – 22.06.2026	виконано
10	Порівняння з існуючими підходами	22.06.2026 – 23.06.2026	виконано
11	Формулювання висновків і напрямів подальшого розвитку	23.06.2026 – 24.06.2026	виконано
12	Написання записки: теоретична частина	25.06.2026 – 26.06.2026	виконано
13	Написання записки: опис програмного застосунку	27.06.2026 – 28.06.2026	виконано
14	Підготовка до нормоконтролю	29.06.2026 – 30.06.2026	виконано
15	Захист перед ЕК	09.07.2026	

Дата видачі завдання 8 червня 20 26 р.

Здобувач _____


(підпис)

Керівник роботи _____

(підпис)

проф. Лариса ЧАЛА

(посада, власне ім'я, прізвище)

РЕФЕРАТ

Пояснювальна записка: 85 с., 11 рис., 2 табл., 1 дод., 89 джерел.

ВИЛУЧЕННЯ ОЗНАК, ГЕНЕРАЦІЯ СИНТЕТИЧНИХ ДАНИХ, ГЛИБОКЕ НАВЧАННЯ, ДАТАСЕТ, ЗГОРТКОВА НЕЙРОННА МЕРЕЖА, МАШИННЕ НАВЧАННЯ, НЕЙРОННА МЕРЕЖА, ПОКАЗНИК ХЕРСТА, ПРОГНОЗУВАННЯ, РОБАСТНІСТЬ, СТАТИСТИЧНІ КРИТЕРІЇ, ЧАСОВІ РЯДИ, ШТУЧНИЙ ІНТЕЛЕКТ.

Об'єкт дослідження – процес оцінювання показника Херста часових рядів із властивостями фрактального броунівського руху.

Предмет дослідження – архітектура згорткової нейронної мережі з шаром глобального пулінгу для оцінювання показника Херста.

Мета роботи – розробка методу оцінювання показника Херста часових рядів будь-якої довжини, симульованих фрактальним броунівським рухом, на основі згорткової нейронної мережі з шаром глобального пулінгу.

Методи дослідження – чисельне моделювання траєкторій фрактального броунівського руху, порівняльний аналіз статистичних методів та методів машинного навчання, обчислювальний експеримент на синтетичних даних, аналіз метрик точності оцінювання.

У роботі запропоновано удосконалення згорткової нейронної мережі, призначеної для оцінювання показника Херста часових рядів із властивостями фрактального броунівського руху, що на відміну від опублікованих раніше підходів використовує коректну нормалізацію вхідного ряду та включає в себе додатковий шар вилучення ознак методом узагальненого показника Херста другого порядку. Запропонований метод суттєво покращує точність оцінювання коротких часових рядів порівняно з опублікованим раніше рішенням.

ABSTRACT

Bachelor's thesis contains: 85 pp., 11 fig., 2 tabl., 1 ann., 89 references.

ARTIFICIAL INTELLIGENCE, CONVOLUTIONAL NEURAL NETWORK, DATASET, DEEP LEARNING, FEATURE EXTRACTION, FORECASTING, HURST EXPONENT, MACHINE LEARNING, NEURAL NETWORK, ROBUSTNESS, STATISTICAL CRITERIA, SYNTHETIC DATA GENERATION, TIME SERIES.

The object of the study is the process of estimating the Hurst exponent of time series with properties of fractional Brownian motion.

The subject of the study is the architecture of a convolutional neural network with a global pooling layer for estimating the Hurst exponent.

The objective of the study is to develop a method for estimating the Hurst exponent of time series of arbitrary length, simulated by fractional Brownian motion, based on a convolutional neural network with a global pooling layer.

The research methods include numerical simulation of fractional Brownian motion trajectories, comparative analysis of statistical and machine learning methods, computational experiments on synthetic data, and analysis of estimation accuracy metrics.

The paper proposes an improvement to a convolutional neural network designed for estimating the Hurst exponent of time series with fractional Brownian motion properties. Unlike previously published approaches, it utilizes proper normalization of the input series and includes an additional feature extraction layer using the second-order generalized Hurst exponent method. The proposed method significantly improves the estimation accuracy for short time series compared to previously published solutions.

ЗМІСТ

Перелік умовних позначень, символів, одиниць, скорочень і термінів	8
Вступ.....	9
1 Аналіз процесу оцінювання показника Херста для часових рядів фрактального броунівського руху	10
1.1 Постановка задачі та аналіз об'єкта дослідження	10
1.1.1 Від класичної до аномальної дифузії.....	11
1.1.2 Фізичний зміст показника Херста та режими руху.....	12
1.2 Актуальність та цільова аудиторія дослідження	16
1.2.1 Фрактальна фармакокінетика та моделювання поведінки ліків	16
1.2.2 Діагностика патологічних станів за фізіологічними сигналами	18
1.2.3 Аналіз спостережень окремих частинок у клітинній біології...	19
1.3 Аналіз існуючих методів оцінювання показника Херста для випадку часових рядів фрактального броунівського руху	20
1.4 Недоліки існуючих рішень та обґрунтування потреби в розробці нових рішень	21
1.5 Формулювання задачі дослідження	22
2 Методи оцінювання показника Херста для фрактального гауссового шуму	25
2.1 Спектральний метод Уїттла	25
2.2 Метод узагальненого показника Херста.....	29
2.3 Метод Колмогорова-Смірнова.....	31
2.4 Метод RANDI на основі рекурентних нейромереж	34
2.5 Гібридний метод CINNAMON.....	35
2.6 Порівняльний аналіз та обґрунтування потреби в новому підході ...	36
2.7 Оцінювання показника Херста на основі згорткових нейронних мереж	38
3 Реалізація оцінювача показника Херста на основі CNN та експериментальні дослідження.....	46

3.1 Реалізація моделі зі згортковою мережею та шаром глобального усередненого пулінгу.....	47
3.2 Експеримент з адаптацією нормування до задачі аналізу часових рядів фрактального броунівського руху.....	54
3.3 Експерименти з ruHurstEE як екстрактором ознак.....	58
3.4 Спрощення архітектури згорткового екстрактора ознак.....	61
3.5 Дослідження робастності моделі до адитивного шуму	64
3.6 Порівняльний експеримент на датасеті AnDi-2.....	67
Висновки	70
Перелік джерел посилання	73
Додаток А Відомість кваліфікаційної роботи.....	85

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

СКВ – середньоквадратичне відхилення;

ШІ – штучний інтелект;

α – показник аномальної дифузії;

CNN – Convolutional Neural Networks – згорткові нейронні мережі;

fBm – fractional Brownian motion – фрактальний броунівський рух;

FFT – Fast Fourier Transform – швидке перетворення Фур'є;

fGn – fractional Gaussian noise – фрактальний гауссів шум;

GAP – Global Average Pooling – шар глобального усередненого пулінгу;

GHE – Generalized Hurst Exponent – узагальнений показник Херста;

H – показник Херста;

HurstEE – нейромережевий шар для оцінювання показника Херста;

LSTM – Long Short-Term Memory – довгострокова короткочасна пам'ять;

MAE – Mean Absolute Error – середня абсолютна похибка;

MSD – Mean Squared Displacement – середнє квадратичне зміщення;

TA-MSD – Time-Averaged Mean Squared Displacement – усереднене за часом середнє квадратичне зміщення.

ВСТУП

Аналіз часових рядів із фрактальними властивостями є однією з актуальних задач сучасної науки про дані. Математичною основою таких процесів слугує фрактальний броунівський рух, ключовим параметром якого є показник Херста H , що характеризує ступінь довготривалої пам'яті та самоподібності процесу. Оцінювання показника Херста знаходить застосування в біофізиці, матеріалознавстві, телекомунікаціях та інших галузях, де спостерігаються процеси з аномальною дифузією.

Практична складність задачі полягає в тому, що реальні часові ряди часто є короткими та зашумленими. Класичні статистичні методи оцінювання показника Херста – такі як метод нормованого розмаху, усереднене за часом середньоквадратичне зміщення та метод детрендованого флуктуаційного аналізу – демонструють значне зміщення оцінок і високу дисперсію на коротких вибірках. Водночас існуючі нейромеревеві підходи, як-от RANDI, є універсальними й не враховують специфіки рядів, симульованих саме фрактальним броунівським рухом.

Метою роботи є розробка методу оцінювання показника Херста часових рядів будь-якої довжини, симульованих фрактальним броунівським рухом, на основі згорткової нейронної мережі з шаром глобального пулінгу.

Об'єктом дослідження є процес оцінювання показника Херста часових рядів із властивостями фрактального броунівського руху. Предметом дослідження є архітектура згорткової нейронної мережі з шаром глобального пулінгу для оцінювання показника Херста.

Результати роботи можуть бути застосовані в задачах аналізу траєкторій біомолекул у клітинній біофізиці, медицині, матеріалознавстві, аналізі часових рядів.

1 АНАЛІЗ ПРОЦЕСУ ОЦІНЮВАННЯ ПОКАЗНИКА ХЕРСТА ДЛЯ ЧАСОВИХ РЯДІВ ФРАКТАЛЬНОГО БРОУНІВСЬКОГО РУХУ

1.1 Постановка задачі та аналіз об'єкта дослідження

Світ навколо нас перебуває в постійному, безперервному русі, природа якого варіюється від суворо детермінованих траєкторій небесних тіл до хаотичних, на перший погляд, блукань мікроскопічних частинок у рідинах та газах. У сучасній науці, зокрема в оптичній мікроскопії, біофізиці, матеріалознавстві, аналіз поведінки окремих агентів або молекул дозволяє не лише визначити кінетику їхньої взаємодії, але й кількісно оцінити фундаментальні динамічні властивості середовища, у якому вони перебувають.

Фундаментальною концепцією для опису таких стохастичних процесів тривалий час залишався класичний броунівський рух – безладний рух дрібних частинок, спричинений тепловими флуктуаціями середовища, який традиційно моделюється за допомогою стандартного броунівського руху (вінерівського процесу) [1].

Однак реальні процеси, що спостерігаються в живих системах та складних фізичних середовищах, нерідко демонструють поведінку, що суттєво відхиляється від класичних припущень – зокрема, наявність довгострокової пам'яті, самоподібності та аномальної дифузії. Саме тут на перший план виходить наука про дані, завданням якої є розроблення та вдосконалення методів кількісного аналізу таких процесів і часових рядів: від статистичного оцінювання параметрів моделей до алгоритмів машинного навчання, здатних виявляти приховані закономірності у великих масивах траєкторних даних і забезпечувати надійну ідентифікацію типу та режиму досліджуваного процесу.

1.1.1 Від класичної до аномальної дифузії

Класична теорія дифузії, заснована на піонерських роботах Альберта Ейнштейна та Маріана Смолуховського початку XX століття, стверджує, що для нормальної дифузії середньоквадратичне зміщення (Mean Squared Displacement, MSD) частинки лінійно залежить від часу. Ця лінійність є наслідком центральної граничної теореми для суми незалежних однаково розподілених випадкових величин. Це означає, що процес не має «пам'яті», а його прирости є незалежними гауссовими величинами, що математично виражається формулою [2]:

$$\text{MSD}(t) \equiv \langle (x(t) - x(0))^2 \rangle = 2dDt, \quad (1.1)$$

де d позначає розмірність простору, у якому відбувається рух;

D є коефіцієнтом дифузії, що характеризує швидкість розповсюдження частинок.

Однак, з розвитком експериментальних методів, дослідники все частіше стикаються із ситуаціями, коли класична теорія виявляється неспроможною описати спостережувані явища. У реальних гетерогенних системах – таких як цитоплазма живих клітин, насичена органелами та цитоскелетом, пористі матеріали зі складною топологією, в'язкопружні полімерні розчини – динаміка частинок часто демонструє значні відхилення від лінійного закону. Таке явище класифікують як аномальну дифузію [3]. Для неї часова залежність MSD набуває загального степеневих вигляду:

$$\text{MSD}(t) \propto K_{\alpha} t^{\alpha}, \quad (1.2)$$

де K_{α} – це узагальнений коефіцієнт дифузії;

α – безрозмірний показник аномальної дифузії. Саме цей показник стає ключовим індикатором, що визначає режим транспорту речовини або інформації в системі.

Ключовою та найбільш універсальною математичною моделлю для опису таких процесів, особливо в контексті біологічних систем, є фрактальний броунівський рух (fractional Brownian motion, fBm) [4]. Це центрований гауссів процес $B_H(t)$ з неперервним часом, який, на відміну від класичного броунівського руху, характеризується стаціонарними, але залежними приростами. Ця модель, запропонована Мандельбротом та Ван Нессом, природним чином враховує властивість самоподібності та наявність довготривалої залежності (пам'яті) в системі. Основною характеристикою цього процесу є показник Херста (H), який безпосередньо пов'язаний з показником аномальної дифузії співвідношенням:

$$\alpha = 2H. \quad (1.3)$$

1.1.2 Фізичний зміст показника Херста та режими руху

Фізична сутність показника Херста розкривається через кореляційні властивості приростів процесу, які формують так званий фрактальний гауссів шум (fractional Gaussian noise, fGn). Автокореляційна функція цього шуму для fBm спадає степеневим чином, що свідчить про повільне затухання кореляцій (довгострокову пам'ять), тобто про те, що минулі події мають суттєвий вплив на майбутню поведінку системи навіть на значних часових інтервалах [5]. Залежно від конкретного значення H , виділяють три фундаментальні режими руху, зображені на рисунку 1.1, кожен з яких відповідає різним фізичним механізмам.

Субдифузія спостерігається при $0 < H < 0.5$ (відповідно, $\alpha < 1$). У цьому випадку процес є антиперсистентним. Це означає, що прирости процесу негативно корельовані: якщо частинка здійснила крок у певному

напрямку, існує висока статистична ймовірність, що наступний крок буде зроблено в протилежному напрямку. У біофізиці такий тип руху часто свідчить про наявність просторових перешкод, ефект «скупчення» молекул, перебування частинки у в'язкопружному середовищі або її тимчасове захоплення в «пастки» [6]. Прикладом може слугувати рух білків у щільній цитоплазмі або латеральна дифузія ліпідів у клітинних мембранах.

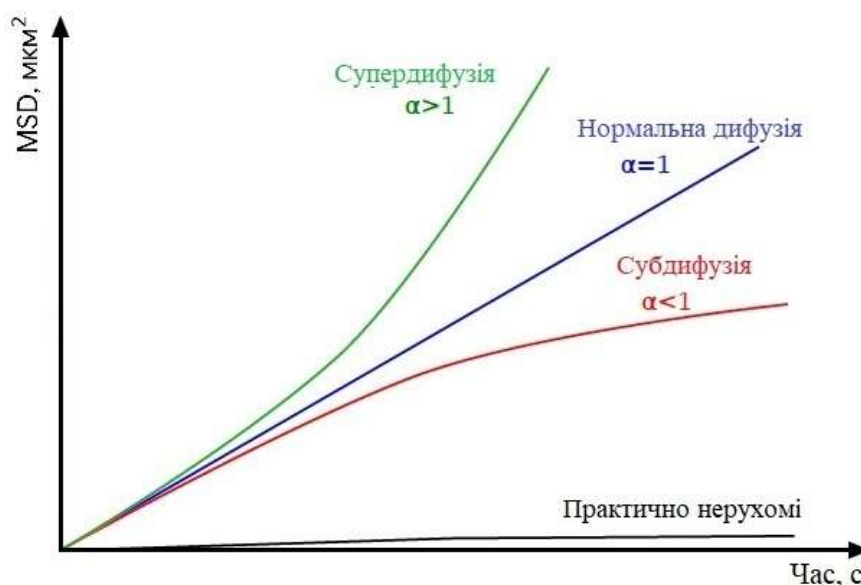


Рисунок 1.1 – Середньоквадратичне зміщення для різних типів дифузії [7]

Нормальна дифузія відповідає значенню $H = 0.5$ ($\alpha = 1$). Це граничний випадок, що відповідає класичному броунівському руху. Прирости є незалежними й нескорельованими (білий гауссів шум), а сам fVtm при $H=0.5$ збігається з класичним вінерівським процесом, який є марковським, тобто таким, що «не має пам'яті». Цей режим розділяє два типи аномальної поведінки й слугує базовою лінією для порівняння.

Супердифузія виникає при $0.5 < H < 1$ ($\alpha > 1$). У цьому стані процес є персистентним. Існує позитивна кореляція між приростами: частинка має виражену тенденцію зберігати напрямок свого руху. У біологічних системах це є характерною ознакою активного внутрішньоклітинного транспорту, коли молекулярні мотори переносять вантажі вздовж мікротрубочок,

витрачаючи енергію АТФ [8]. В економіці такий режим відповідає трендовим процесам на фінансових ринках, де зростання ціни активу підвищує ймовірність подальшого зростання [9].

На рисунку 1.1 криволінійний характер залежностей MSD у лінійних координатах ускладнює кількісне розрізнення режимів дифузії, особливо коли значення α близькі до одиниці. Цю проблему усуває перехід до подвійних логарифмічних координат (рисунок 1.2).

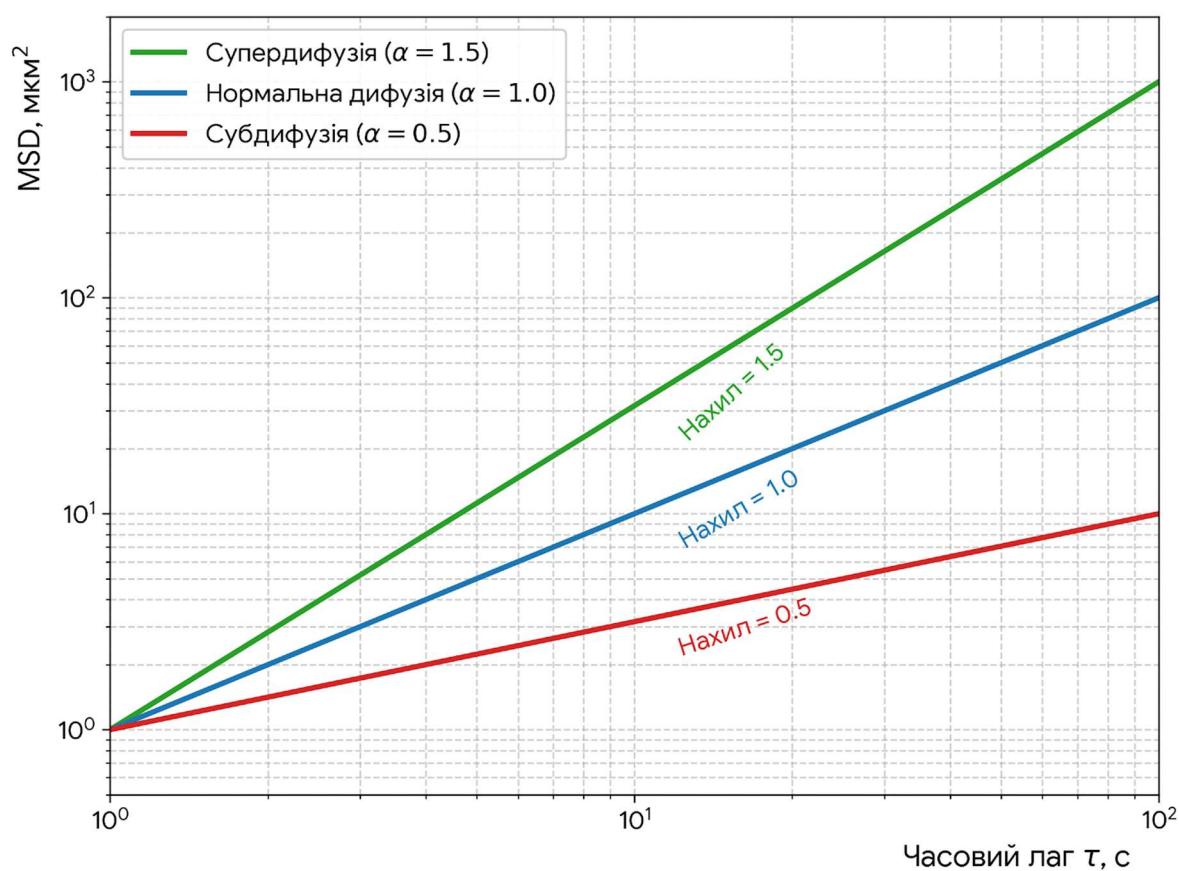


Рисунок 1.2 – Середньоквадратичне зміщення для різних типів дифузії у подвійних логарифмічних координатах

Степенева залежність (1.2) після логарифмування обох частин набуває вигляду:

$$\log \text{MSD}(t) = \alpha \cdot \log(t) + \log(K_\alpha), \quad (1.4)$$

тобто є рівнянням прямої в координатах ($\log t$, $\log \text{MSD}$), у якій кутовий коефіцієнт безпосередньо дорівнює показнику аномальної дифузії α , а вільний член визначається узагальненим коефіцієнтом дифузії K_α , введеним у формулі (1.2). На рисунку 1.2 проілюстровано асимптотичну поведінку MSD у подвійних логарифмічних координатах для трьох режимів – субдифузії, нормальної дифузії та супердифузії, – представлених прямими з нахилами 0.5, 1.0 та 1.5 відповідно. Слід зауважити, що ідеально лінійний характер цих залежностей реалізується лише в асимптотичному режимі, тобто за умови достатньо довгих часових рядів; для скінченних реалізацій спостерігаються відхилення від прямої, особливо на малих та великих часових лагах τ .

Ця властивість лежить в основі одного з найпоширеніших методів оцінки показника аномальної дифузії – аналізу часово-усередненого середньоквадратичного зміщення (Time-Averaged MSD, TA-MSD). За емпіричною траєкторією частинки обчислюють значення TA-MSD для дискретного набору часових лагів τ , наносять отримані точки в лог-лог координатах і підганяють до них пряму методом найменших квадратів. Кутовий коефіцієнт регресійної прямої β є оцінкою показника α , а з огляду на співвідношення (1.3) між показником аномальної дифузії та показником Херста – також і оцінкою останнього: $\hat{H} = \alpha/2$. Простота, обчислювальна ефективність та прозора фізична інтерпретація результату зробили цей підхід де-факто стандартом первинного аналізу одночастинкових траєкторій і базовою лінією для порівняння зі складнішими методами оцінки H , які детально розглядаються в розділі 2.

Варто зазначити, що хоча fBm є найбільш поширеною моделлю, сучасні дослідження, зокрема в рамках міжнародних змагань AnDi Challenge, розглядають й інші моделі аномальної дифузії, такі як блукання з неперервним часом (Continuous Time Random Walk, CTRW) або масштабований броунівський рух (Scaled Brownian Motion, SBM) [10], [11]. Ці моделі можуть поводити себе складніше, наприклад, коли середні

значення, розраховані для однієї довгої траєкторії, не збігаються з середніми, розрахованими по групі багатьох коротких траєкторій (порушення ергодичності). Крім того, для таких процесів може не виконуватися класичне співвідношення $H = \alpha / 2$. Проте фрактальний гауссів шум fGn, який є приростами fBm, залишається класичною гауссовою моделлю, яка коректно описує стаціонарні процеси з довготривалою пам'яттю, що робить точну оцінку H критично важливою й першочерговою задачею для ідентифікації типу руху. Нещодавні теоретичні роботи, зокрема [12], навіть розглядають узагальнення цієї моделі, такі як фрактальний броунівський рух Рімана-Ліувілля з випадковим показником Херста, що підкреслює надзвичайну складність стохастичної природи реальних фізичних та біологічних систем.

1.2 Актуальність та цільова аудиторія дослідження

Проблема точної оцінки показника Херста виходить далеко за межі суто теоретичної фізики або математичної статистики. Вона є критично важливим інструментом у широкому спектрі прикладних наук, де здатність кількісно охарактеризувати ступінь хаотичності, персистентності або детермінованості процесу відкриває принципово нові можливості для прогнозування, діагностики та оптимізації. Аналіз сучасної наукової літератури дозволяє виділити декілька ключових сфер, де використання моделей фрактального броунівського руху та точне визначення H є найбільш ефективним та затребуваним.

1.2.1 Фрактальна фармакокінетика та моделювання поведінки ліків

Фармакокінетика є розділом фармакології, який вивчає кількісні закономірності руху лікарських речовин в організмі. Вона описує, що відбувається з препаратом з моменту його введення і до моменту повного

виведення. Класично цей процес охоплює чотири основні етапи, які часто скорочують аббревіатурою ADME: Absorption (всмоктування), Distribution (розподіл), Metabolism (метаболізм, або біотрансформація) та Excretion (виведення) [13].

Мета фармакокінетики – побудова математичних моделей, які дозволяють прогнозувати концентрацію препарату в крові та тканинах у будь-який момент часу, що є критично важливим для визначення оптимального режиму дозування. Традиційна фармакокінетика протягом десятиліть базується на класичних компартментних моделях, що спираються на припущення про миттєве перемішування речовини та гомогенність середовища, фактично використовуючи закон дії мас для ідеальних розчинів.

Однак, як слушно зазначають дослідники в [14], [15], живий організм є складним гетерогенним середовищем із яскраво вираженою ієрархічною та фрактальною структурою. Це стосується архітектури розгалуження кровоносної системи, губчастої структури легеневої тканини, фрактальної шорсткості клітинних мембран та нерівномірного розподілу рецепторів на них.

У таких складних умовах процеси всмоктування (абсорбції), розподілу в тканинах та виведення ліків часто суттєво відхиляються від класичної експоненційної кінетики першого порядку. Використання моделей фрактального броунівського руху дозволяє моделювати ці складні явища значно точніше. За таких умов даними для задачі будуть часові ряди концентрації лікарського засобу в плазмі крові пацієнта або дані про розподіл міченої речовини в тканинах, отримані методами ПЕТ/СТ.

Насамперед, розраховане значення H використовується для уточнення моделей розподілу ліків. Якщо $H < 0.5$ (субдифузія), це свідчить про затримку препарату в тканинах. У [16] запропоновано спеціалізовану модель «Pathwise Fractional one Compartment Intra-Veinous Bolus Model», яка використовує H для корекції рівнянь. Це дозволяє лікарям точніше

прогнозувати пікові концентрації препарату та час їх досягнення, уникаючи передозування.

Крім того, показник Херста є критичним параметром для прогнозування токсичності. Низьке значення H вказує на ймовірність «застрягання» молекул у певних компартментах (ефект пасток), що може призводити до локального накопичення токсинів. Наприклад, дослідження фармакокінетики таких препаратів, як аміодарон (потужний антиаритмічний засіб), показали, що їх поведінка краще описується стохастичними моделями з урахуванням аномальної дифузії [15]. Оцінка H дозволяє фармацевтам передбачити ризики накопичення препарату при тривалому прийомі.

1.2.2 Діагностика патологічних станів за фізіологічними сигналами

Надзвичайно перспективною сферою застосування розроблених методів є біомедична інженерія та функціональна діагностика. Фізіологічні сигнали, такі як електрична активність мозку (ЕЕГ), варіабельність серцевого ритму (ВСР) або патерни ходи, є результатом складної взаємодії багатьох регуляторних механізмів організму. Як показали дослідження [17], ці сигнали мають фрактальну структуру, а показник Херста є надійним кількісним індикатором їхньої складності та «здоров'я» системи. За таких умов на вхід нашої моделі будуть подані часові ряди фізіологічних сигналів (інтервали R-R з ЕКГ, амплітуда сигналу ЕЕГ), записані протягом короткого проміжку часу (наприклад, під час стандартного обстеження).

Оцінка H дозволяє лікарям виявляти ранні ознаки нейродегенеративних захворювань. Наприклад, при хворобі Альцгеймера або Паркінсона спостерігається зниження фрактальної складності сигналів ЕЕГ, що проявляється в зміні показника Херста ще до появи явних клінічних симптомів [18]. Модель, здатна працювати з короткими часовими рядами, може бути інтегрована в портативні діагностичні пристрої для скринінгу.

Також це критично для моніторингу серцево-судинної системи. Здоровий серцевий ритм характеризується певною мірою хаотичності (H близьке до 1), що свідчить про високу адаптивність системи. При серцевій недостатності або після інфаркту ритм стає більш регулярним або некорельованим, що миттєво відображається в зміні H . Використання точних методів оцінювання H дозволяє кардіологам у реальному часі відстежувати стан пацієнта в реанімації та прогнозувати ризики раптової зупинки серця [19].

1.2.3 Аналіз спостережень окремих частинок у клітинній біології

Окрім медичних застосувань, оцінка H фактично стала стандартом у фундаментальній клітинній біології. Розвиток методів флуоресцентної мікроскопії високої роздільної здатності дозволяє дослідникам відстежувати траєкторії окремих молекул (білків, мРНК, вірусів, органел) всередині живих клітин з нанометровою точністю. Вхідними даними в цьому контексті є саме ці траєкторії – часові ряди координат (x , y , z) частинок, отримані методом відстеження окремих частинок.

Як показано в численних роботах, що лягли в основу міжнародних змагань AnDi Challenge [10], [11], траєкторії, наприклад, теломер у ядрі клітини або рецепторів на плазматичній мембрані, є яскравими прикладами аномальної дифузії. Значення показника Херста тут виступає в ролі чутливого біомаркера: його зміна може свідчити про функціональну активацію клітини, зміну мікров'язкості цитоплазми, перебудову цитоскелета або початок патологічних процесів (наприклад, вірусної інфекції або апоптозу). Таким чином, оцінене значення H дозволяє біологам кількісно відрізняти пасивний рух від активного транспорту [8], ідентифікувати залучення молекул у структури цитоскелета та визначати ступінь обмеженості руху.

1.3 Аналіз існуючих методів оцінювання показника Херста для випадку часових рядів фрактального броунівського руху

Попри наявність глибоко розробленої теоретичної бази, практична оцінка показника Херста залишається складною та нетривіальною обчислювальною задачею. Проблему ускладнює та обставина, що різні методи мають різні припущення щодо природи процесу, потребують різного обсягу даних та демонструють суттєво відмінну поведінку в умовах шуму й кінцевих вибірок. Існуючі методи можна класифікувати на три великі групи залежно від підходу до обробки даних [20].

До першої групи належать часові методи (Time-domain methods), які працюють безпосередньо з часовим рядом та його статистичними характеристиками в часовій області. Найвідомішим є R/S аналіз (Rescaled Range Analysis) – класичний метод, запропонований самим Херстом, що базується на аналізі масштабування нормованого розмаху накопиченого відхилення від середнього. Широко використовується також DFA (Detrended Fluctuation Analysis) [21], який ефективно усуває нестационарні тренди через локальну апроксимацію поліномами, та його мультифрактальне узагальнення MF-DFA. TA-MSD є базовим методом у біофізиці, що оцінює α через нахил log-log графіка усередненого за часом MSD, з якого для fBm обчислюється $H = \alpha/2$. Метод GHE (Generalized Hurst Exponent) [22] використовує масштабування моментів q-го порядку і дозволяє діагностувати мультифрактальність. Метод Higuchi оцінює H через фрактальну розмірність часового ряду. Окрему підгрупу становлять геометричні та статистичні методи: Total Triangle Areas (TTA) і його вдосконалена версія TA, метод Колмогорова-Смірнова (KS), який порівнює емпіричні розподіли характеристик процесу при різних значеннях H [23], а також байєсівські підходи LSV та LSSD.

Друга група – це спектральні методи (Spectrum-domain methods), які використовують частотний аналіз. Оцінювач Уїттла є параметричним

методом максимальної правдоподібності в частотній області, запропонованим у 1953 році й адаптованим для довгопам'ятних процесів [24]. Він базується на мінімізації відхилення між емпіричною періодограмою та теоретичним спектром fGn. Напівпараметричний локальний оцінювач Уіттла обмежує аналіз низькочастотною областю спектра, що підвищує робастність до відхилень від ідеальної моделі. Також застосовуються методи на основі вейвлет-перетворення, які аналізують масштабувальні властивості сигналу одночасно на декількох часових масштабах [25], [26], [27].

Третя група – це сучасні методи на основі машинного навчання та глибокого навчання, які відкривають принципово нові можливості для оцінювання H [28], [29], [30]. Метод RANDI використовує рекурентні нейронні мережі LSTM, навчені на синтетичних траєкторіях п'яти моделей аномальної дифузії, для одночасного вирішення задач класифікації, регресії та сегментації [31]. CINNAMON поєднує глибокі нейронні мережі, градієнтний бустинг на рукотворних статистичних ознаках та аналітичні оцінювачі для аналізу неоднорідних траєкторій [32]. HurstEE реалізує диференційований шар для оцінювання H усередині глибоких нейромереж [33]. Ці методи, як правило, навчаються на великих наборах синтетичних даних fGn, щоб автоматично вивчити складні статистичні патерни, недоступні для класичних підходів, що базуються на аналітичних формулах.

1.4 Недоліки існуючих рішень та обґрунтування потреби в розробці нових рішень

Як показали результати міжнародних змагань Anomalous Diffusion (AnDi) Challenge, більшість класичних статистичних методів мають суттєві, іноді критичні, обмеження при роботі з даними, що імітують реальні експериментальні умови [10]. Порівняльні дослідження методів

оцінювання показника Херста виявили ряд системних проблем, властивих як традиційним підходам, так і існуючим нейромережевим рішенням.

По-перше, існує критична чутливість до довжини ряду. Траєкторії відстеження окремих частинок часто містять лише 20–50 точок через технічні обмеження: фотовигорання флуорофорів, вихід частинки з фокальної площини або розпад молекулярного комплексу. На таких малих вибірках асимптотичні властивості класичних методів – R/S, DFA, GHE, оцінювача Уїттла – ще не встигають проявитися, а статистичні похибки стають неприйнятно великими [34].

По-друге, значним фактором є вплив адитивного шуму вимірювання. Експериментальні дані завжди містять похибки локалізації центру мас частинки, які накладаються на істинний рух. Цей шум спотворює кореляційну структуру fBm особливо на малих часових масштабах: GHE та оцінювач Уїттла систематично занижують H , оскільки шум ефективно маскує довготривалі кореляції. Метод KS частково стійкіший до шуму завдяки порівнянню цілих розподілів, однак і він вимагає калібрування при значному рівні локалізаційного шуму [23].

По-третє, суттєвою проблемою є фіксованість архітектури існуючих нейромережових рішень щодо довжини входу. RANDI вирішує цю проблему частково – через навчання окремих мереж для різних діапазонів довжин і інтерполяцію між ними [31], що є технічно складним рішенням. RANDI та подібні моделі часто вимагають вхідних даних фіксованої довжини, що призводить до потреби в обрізанні, яке спотворює статистичні властивості траєкторій. Ідеальна архітектура має нативно підтримувати вхід змінної довжини.

1.5 Формулювання задачі дослідження

Метою роботи є розробка методу оцінювання показника Херста часових рядів будь-якої довжини, симульованих фрактальним броунівським

рухом, на основі згорткової нейронної мережі з шаром глобального пулінгу. Для досягнення поставленої мети необхідно вирішити такі завдання:

- проаналізувати існуючі статистичні й нейромережеві методи оцінювання показника Херста та виявити їх недоліки для коротких часових рядів;
- розробити пайплайн для генерації синтетичних даних на основі траєкторій фрактального броунівського руху з відомими значеннями H ;
- реалізувати архітектуру згорткової нейронної мережі з шаром глобального пулінгу, здатної обробляти ряди довільної довжини;
- провести порівняльний експериментальний аналіз методу та його запропонованих удосконалень з існуючими підходами на синтетичних даних та датасеті аномальної дифузії AnDi.

Вхідними даними є часові ряди приростів фрактального гауссового шуму. Використання приростів замість абсолютних координат дозволяє працювати зі стаціонарним процесом, що спрощує навчання мережі. Довжина рядів може бути довільною (починаючи від кількох точок), а самі дані можуть містити адитивний гауссів шум, що імітує похибки вимірювання. Вихідними даними системи мають бути оцінки – скалярне значення показника Херста $H \in (0, 1)$.

Ключовими вимогами до розроблюваного рішення є гнучкість, що передбачає архітектуру з механізмами (зокрема, шар Global Average Pooling), які дозволяють обробляти вхідні дані змінної довжини без необхідності повторного навчання моделі або попередньої обробки даних (обрізання/паддінг). Також важливою є точність, адже модель повинна демонструвати показники, що перевершують результати класичних статистичних методів (таких як TA-MSD, DFA) на коротких траєкторіях (<50 точок). Нарешті, робастність вимагає забезпечення стійкості оцінки до локалізаційного шуму різної інтенсивності.

Таким чином, у роботі було здійснено комплексний аналіз задачі оцінювання показника Херста для часових рядів, що моделюються фрактальним броунівським рухом.

Установлено, що фрактальний броунівський рух є однією з поширених базових математичних моделей для опису стохастичних процесів у складних гетерогенних середовищах – від біологічних тканин на мікрорівні до комплексних природничих систем. Показник Херста при цьому виступає фундаментальним індикатором, що визначає тип дифузії (суб- чи супердифузія), ступінь персистентності процесу та наявність довготривалої пам'яті в системі.

Визначено, що найбільш перспективними та соціально значущими сферами практичного застосування цих оцінок є фрактальна фармакокінетика та діагностика патологічних станів за фізіологічними сигналами. Врахування показника Херста при моделюванні розподілу ліків дозволяє прогнозувати їх концентрацію в тканинах та органах значно точніше, ніж це роблять класичні моделі. У сфері біомедичної діагностики точна оцінка фрактальних властивостей сигналів дозволяє виявляти ранні ознаки нейродегенеративних та серцево-судинних захворювань, що має прямий вплив на якість життя пацієнтів.

Виявлено, що існуючі статистичні методи (як часові, так і спектральні) є недостатньо ефективними для умов сучасного експерименту, які характеризуються короткими та зашумленими траєкторіями. Це обґрунтовує необхідність розробки та дослідження спеціалізованих нейромережових архітектур, зокрема на базі CNN з використанням шарів глобального пулінгу, які здатні забезпечити стійку оцінку параметрів аномальної дифузії без необхідності ручного налаштування під кожен окремий експеримент.

2 МЕТОДИ ОЦІНЮВАННЯ ПОКАЗНИКА ХЕРСТА ДЛЯ ФРАКТАЛЬНОГО ГАУССОВОГО ШУМУ

Будь-яке дослідження, спрямоване на розробку нових алгоритмів обробки стохастичних часових рядів, потребує глибокого розуміння існуючих методологічних підходів, їхніх теоретичних засад та практичних обмежень. У контексті задачі оцінювання показника Херста, який виступає ключовим індикатором довгострокової пам'яті та самоподібності процесу, особливого значення набуває аналіз методів, що продемонстрували найкращі результати в сучасних порівняльних дослідженнях. Для забезпечення об'єктивності порівняльного аналізу та обґрунтування доцільності розробки нових підходів, у даному розділі детально розглядаються п'ять методів, обраних за критерієм ефективності та різноманітності підходів: оцінювач Уїтгла як представник спектральних методів, узагальнений показник Херста з групи часових методів, метод Колмогорова-Смірнова як непараметричний статистичний підхід, а також два сучасних методи на основі машинного навчання – CINNAMON та RANDI. Такий вибір дозволяє охопити три принципово різні парадигми обробки даних та виявити їхні характерні переваги й недоліки.

2.1 Спектральний метод Уїтгла

Теоретичні засади спектрального підходу до оцінювання показника Херста базуються на фундаментальній властивості фрактальних процесів – специфічній формі спектральної густини потужності, що підпорядковується степеневому закону. Ця властивість є безпосереднім наслідком самоподібної природи fBm та довготривалої кореляційної структури його приростів. Ідею максимально правдоподібного оцінювання в частотній області вперше систематично виклав Уїтгл у 1953 році [35].

Спектральна густина потужності фрактального гауссового шуму $S(f)$ на частоті f описується степеневим законом

$$S(f) \propto f^{(-\beta)}, \quad (2.1)$$

де спектральний показник β безпосередньо пов'язаний з показником Херста співвідношенням $\beta = 2H - 1$. Ця залежність демонструє, що для персистентних процесів з $H > 0.5$ спектр характеризується переважанням низькочастотних компонент, що відповідає наявності довготривалих кореляцій у системі. Навпаки, для антиперсистентних процесів з $H < 0.5$ спостерігається відносне посилення високочастотних компонент, що свідчить про тенденцію до швидких осциляцій навколо середнього значення [36].

Оцінювач Уїтла є параметричним методом максимальної правдоподібності в частотній області. На відміну від непараметричних підходів, він базується на явній параметризації спектральної густини потужності через показник Херста H та параметр масштабу σ^2 . Центральним об'єктом є функція правдоподібності в частотній області, яку Уїтл довів, що вона є асимптотично еквівалентною повній гауссовій функції правдоподібності для достатньо довгих рядів [37]. Практично це означає: замість роботи з $N \times N$ коваріаційною матрицею (що є обчислювально надзвичайно дорогим), достатньо порівняти емпіричну та теоретичну спектральні густини на частотах Фур'є.

Центральним об'єктом методу є функція Уїтла – наближена логарифмічна функція правдоподібності в частотній області. Вона будується для ряду $X_1, X_2, \dots, X_N$ на основі емпіричної періодограми $I(f_j)$ та теоретичної спектральної густини $S(f_j; H)$. Мінімізація функції Уїтла по H еквівалентна максимізації наближеної гауссової правдоподібності [38]. Ключова обчислювальна перевага – завдяки FFT складність алгоритму є $O(N \log N)$ замість $O(N^3)$ для прямого методу MLE.

Логарифмічна функція правдоподібності Уїттла має вигляд суми по всіх частотах Фур'є, де кожен доданок включає відношення періодограми до теоретичного спектра та логарифм теоретичного спектра. Мінімізація цієї функції по параметрах моделі еквівалентна максимізації наближеної гауссової функції правдоподібності в частотній області. Для фрактального гауссового шуму вектор параметрів включає показник Херста H та параметр масштабу σ^2 , причому оцінка H є основним результатом, що цікавить дослідників [39].

Оцінювач Уїттла має ряд важливих статистичних властивостей. За умови коректної специфікації моделі оцінки є асимптотично нормальними та асимптотично ефективними – досягають межі Крамера-Рао для асимптотичної дисперсії [40]. Це означає, що при $N \rightarrow \infty$ оцінювач Уїттла забезпечує найменшу можливу дисперсію серед усіх регулярних незміщених оцінювачів. Крім того, оцінювач дозволяє природним чином обчислити стандартні похибки через обернену матрицю Фішера, що є важливою практичною перевагою порівняно з більшістю інших методів.

Незважаючи на теоретичну оптимальність, практичне застосування класичного оцінювача Уїттла пов'язане з рядом суттєвих обмежень. По-перше, метод є повністю параметричним і вимагає точної специфікації форми спектральної густини потужності на всіх частотах. Якщо реальний процес не є точно фрактальним гауссовим шумом, а лише апроксимується ним, порушення специфікації на високих частотах може призводити до зміщення оцінок [41]. По-друге, періодограма є асимптотично незміщеною, але неконсистентною оцінкою спектра, що проявляється в значній варіабельності оцінок на скінченних вибірках.

Особливо критичним недоліком для біофізичних застосувань є чутливість методу до довжини часового ряду. Як показали дослідження [42], стандартні методи оцінювання показника Херста, включаючи спектральні підходи, демонструють суттєве погіршення точності при роботі з короткими рядами, що містять менше 100 спостережень. Для траєкторій відстеження

окремих частинок, які типово складаються з 20-50 точок, асимптотичні властивості оцінювача Уїттла ще не встигають проявитися, а статистична невизначеність стає неприйнятно великою.

Для подолання проблеми чутливості до специфікації моделі було розроблено напівпараметричний варіант – локальний оцінювач Уїттла. Ідея полягає у використанні лише низькочастотної частини спектра, де степеневий закон виконується найточніше, ігноруючи високочастотну область, де можуть проявлятися відхилення від ідеальної моделі [43]. Функція правдоподібності обчислюється лише для частот, що не перевищують певного порогу, який визначається параметром смуги пропускання m .

Локальний метод Уїттла демонструє кращу робастність до порушень моделі та зберігає консистентність навіть при помірних відхиленнях від фрактальної структури на високих частотах. Водночас, обмеження аналізу лише низькочастотною областю призводить до втрати частини інформації, що міститься в даних, та збільшує дисперсію оцінок порівняно з повним параметричним підходом [43].

Важливим практичним аспектом є вибір оптимального значення параметра смуги пропускання m . Занадто мале значення призводить до великої дисперсії через обмежену кількість частотних точок, що використовуються. Занадто велике значення включає високочастотну область, де степеневий закон може порушуватися, що призводить до зміщення. Оптимальний вибір залежить від невідомих характеристик даних, що створює додаткову невизначеність при практичному застосуванні [44].

Підсумовуючи аналіз оцінювача Уїттла, можна виділити його ключові переваги: строга теоретична обґрунтованість, асимптотична ефективність за умови правильної специфікації моделі, та можливість отримання стандартних похибок оцінок. Водночас, суттєвими недоліками є чутливість до порушень модельних припущень та обчислювальна складність.

2.2 Метод узагальненого показника Херста

Метод узагальненого показника Херста (Generalized Hurst Exponent, GHE) є розширенням класичного підходу до аналізу масштабування на статистичні моменти довільного порядку q [22]. Ключовою теоретичною основою є рівність у розподілі для самоподібних процесів зі стаціонарними приростами, з якої через момент q -го порядку отримується масштабувальний закон $K_q(\tau) \propto \tau^{qH}$. Метод KS-GHE [23] використовує цю рівність у розподілі повністю – мінімізуючи суму KS-відстаней між емпіричними розподілами нормованих приростів для різних τ , що є строгішою умовою.

Теоретичною основою методу є властивість самоафінності фрактальних процесів, згідно з якою статистичні характеристики приростів на різних часових масштабах пов'язані степеневими законами. Для часового ряду $X(t)$ визначається структурна функція q -го порядку $K_q(\tau)$, яка представляє собою усереднений абсолютний момент приростів порядку q для фіксованого часового лагу τ . Математично це записується як усереднення по всіх можливих парах точок з фіксованим лагом величини [22], [45].

$$K_q(\tau) = \langle |X(t + \tau) - X(t)|^q \rangle. \quad (2.2)$$

Для самоафінних процесів структурна функція підпорядковується степеневому закону масштабування

$$K_q(\tau) \propto \tau^{qH(q)}, \quad (2.3)$$

де $H(q)$ є узагальненим показником Херста, який у загальному випадку може залежати від порядку моменту q . Для монофрактальних процесів, до яких належить фрактальний броунівський рух, узагальнений показник $H(q)$ не

залежить від q і збігається зі стандартним показником Херста H для всіх значень q . Ця властивість є критичною для ідентифікації типу процесу та валідації результатів [46].

Оцінювання показника Херста в методі GHE здійснюється через аналіз нахилу залежності логарифма структурної функції від логарифма часового лагу. Побудова графіка $\ln(Kq(\tau))$ від $\ln(\tau)$ для різних значень лагу τ дозволяє візуально оцінити лінійність залежності та визначити діапазон масштабів, на якому виконується степеневий закон. Лінійна регресія в подвійних логарифмічних координатах дає оцінку добутку qH , звідки безпосередньо обчислюється показник Херста [47].

У практичних застосуваннях використовується варіант GHE з $q = 2$, який відповідає аналізу дисперсії приростів. Цей вибір забезпечує найкращий баланс між статистичною ефективністю та стійкістю до викидів. При $q = 1$ структурна функція є менш чутливою до екстремальних значень, але має більшу дисперсію оцінки [48]. При $q > 2$ підвищується чутливість до хвостів розподілу, що може бути корисним для виявлення важких хвостів, але збільшує вплив викидів на результат. Важливою діагностичною можливістю методу GHE є здатність виявляти мультифрактальну структуру процесу [49]. Якщо при різних значеннях q отримуються суттєво різні оцінки $H(q)$, це свідчить про те, що процес не є монофрактальним fBm , а має більш складну природу з різними властивостями масштабування для флуктуацій різної амплітуди. Графік залежності $H(q)$ від q , відомий як спектр узагальненого показника Херста, дозволяє кількісно охарактеризувати ступінь мультифрактальності.

Серед переваг методу GHE слід відзначити його концептуальну простоту та інтуїтивну інтерпретацію результатів. Метод не вимагає складної попередньої обробки даних та може бути застосований безпосередньо до вихідного часового ряду. Можливість варіювання порядку моменту q забезпечує гнучкість підходу та дозволяє адаптувати аналіз до специфіки конкретних даних. Крім того, GHE демонструє більшу стійкість

до одиничних викидів порівняно з методами, що базуються на дисперсії [50].

Водночас, метод GHE має суттєві обмеження, які необхідно враховувати при практичному застосуванні. Подібно до інших часових методів, GHE є чутливим до довжини часового ряду. При малій кількості спостережень на великих лагах статистика структурної функції стає ненадійною, оскільки кількість пар точок з великим лагом стрімко зменшується [51]. Це обмежує діапазон масштабів, на якому можна оцінювати показник, та збільшує невизначеність оцінки.

Серйозною проблемою є вплив адитивного шуму вимірювання на результати аналізу. Шум спотворює кореляційну структуру процесу, особливо на малих часових масштабах, де його внесок є відносно найбільшим. Для GHE це проявляється у викривленні нахилу залежності $\ln(Kq)$ від $\ln(\tau)$ на початковій ділянці графіка, що може призводити до систематичного заниження оцінки H при використанні всього діапазону лагів [52].

Додатковим обмеженням є необхідність вибору діапазону лагів τ для проведення регресії. Занадто малі лаги можуть бути спотворені шумом, занадто великі – мати недостатню статистику. Оптимальний вибір діапазону залежить від характеристик даних та потребує певної експертної оцінки або застосування адаптивних процедур, що ускладнює автоматизацію аналізу [53].

Таким чином, метод узагальненого показника Херста є інструментом для аналізу масштабовувальних властивостей часових рядів з можливістю діагностики мультифрактальності.

2.3 Метод Колмогорова-Смірнова

Метод Колмогорова-Смірнова (KS) для оцінювання показника Херста, запропонований у [23], є непараметричним підходом, що

використовує повну інформацію про форму розподілів, а не лише моменти розподілів (середнє, дисперсію тощо). Теоретичне підґрунтя методу – властивість самоподібних процесів зі стаціонарними приростами: якщо H є справжнім показником Херста, то $F\tau(\tau^H x) = F\tau(x) = F_1(x)$, де $F\tau$ – розподіл $|X(t+\tau) - X(t)|$ [54]. Інакше кажучи, нормовані вибірки $|X(t+\tau) - X(t)| / \tau^H$ повинні мати однаковий розподіл для всіх τ . Статистика Колмогорова-Смірнова D_n визначається як супремум абсолютної різниці між двома емпіричними функціями розподілу. Оцінка H виходить як значення, що мінімізує суму $g(H) = \sum \tau D(F\tau(\tau^H X), F_1(X))$ по всіх обраних лагах τ [55].

Алгоритм KS методу включає шість кроків [23]:

- вибір діапазону лагів τ у вигляді степенів двійки;
- обчислення вибірок нормованих приростів $|X(t+\tau) - X(t)| / \tau^H$ для кожного τ ;
- побудова емпіричних функцій розподілу;
- обчислення KS-відстаней між розподілом для $\tau=1$ та іншими;
- сума відстаней утворює функцію $g(H)$;
- мінімізація $g(H)$ дає оцінку H .

Процедура оцінювання полягає в мінімізації статистики Колмогорова-Смірнова по всіх можливих значеннях показника Херста. Для кожного кандидатного значення $H \in (0, 1)$ обчислюється теоретичний розподіл обраної характеристики, будується емпіричний розподіл за даними та розраховується статистика D_n . Значення H , яке мінімізує D_n , приймається як оцінка показника Херста. Ця оптимізаційна процедура може бути реалізована через пошук на сітці значень H або за допомогою градієнтних методів оптимізації [56], [57].

Важливою перевагою методу Колмогорова-Смірнова є його непараметрична природа, яка не вимагає припущень щодо конкретної форми розподілу похибок або специфіки шуму. Метод є стійким до помірних відхилень від нормальності та демонструє стійкість до присутності важких хвостів у розподілі. Це робить його особливо

придатним для аналізу реальних експериментальних даних, де точна статистична модель може бути невідомою [58].

Серед недоліків методу слід відзначити його обчислювальну складність. Мінімізація статистики по неперервному параметру N вимагає багаторазового обчислення теоретичних розподілів та порівняння їх з емпіричними даними. Якщо теоретичні розподіли не мають аналітичного вигляду й повинні оцінюватися методом Монте-Карло, обчислювальні витрати суттєво зростають [59].

Додатковим обмеженням є чутливість до вибору характеристики процесу, розподіл якої аналізується. Різні функціонали можуть давати різну інформативність щодо показника Херста, і оптимальний вибір не є очевидним. Крім того, для деяких характеристик теоретичні розподіли при різних N можуть бути дуже близькими, що ускладнює розрізнення та збільшує невизначеність оцінки [60].

Важливим практичним аспектом є необхідність калібрування методу. Оскільки теоретичні розподіли залежать не лише від N , але й від інших параметрів процесу (зокрема, від параметра масштабу), ці параметри повинні бути або відомими, або оцінюватися одночасно з показником Херста. Це ускладнює процедуру оцінювання та може вносити додаткову невизначеність [61].

Таким чином, метод Колмогорова-Смірнова представляє альтернативний підхід до оцінювання показника Херста, який демонструє робастність до відхилень від модельних припущень та стабільну продуктивність на гетерогенних даних. Однак обчислювальна складність та необхідність калібрування обмежують його практичне застосування в умовах, коли потрібна швидка автоматизована обробка великих масивів траєкторій.

2.4 Метод RANDI на основі рекурентних неймереж

Розвиток методів глибокого навчання відкрив принципово нові можливості для аналізу траєкторій аномальної дифузії, дозволяючи автоматично виявляти складні патерни в даних без явної специфікації аналітичних залежностей. Метод RANDI (Recurrent Anomalous diffusioN Detection and Inference) є підходом на основі рекурентних нейронних мереж. Його було розроблено та протестовано в рамках міжнародного конкурсу з аномальної дифузії AnDi Challenge у 2020 році, де він продемонстрував найвищу точність класифікації серед 14 поданих методів і входив до трійки кращих методів за всіма дев'ятьма запропонованими задачами – єдиний метод із такою стабільністю результатів [31].

Архітектурною основою всіх мереж RANDI є два послідовних шари довгострокової короткочасної пам'яті LSTM з розмірностями 250 і 50 відповідно. Для задач класифікації за другим LSTM-шаром слідує повнозв'язний шар із 20 вузлами та вихідний шар із 5 вузлами з softmax-активацією, де кожен вузол відповідає ймовірності належності траєкторії до конкретної моделі дифузії. Для задач оцінювання (регресії) вихід другого LSTM-шару подається безпосередньо на вихідний шар із одним вузлом. Функцією втрат для задач регресії є середньоквадратична похибка, а для задач класифікації – категоріальна крос-ентропія; оптимізатором у всіх випадках є Adam [31].

Вхідними даними для мережі є не самі значення траєкторії, а її прирости нормовані так, щоб для кожної траєкторії виконувалося $\Delta\bar{x} = 0$ та $\sigma^2(\Delta x) = 1$. Для підвищення ефективності навчання прирости групуються в блоки фіксованого розміру, що залежить від довжини траєкторії та її просторової розмірності [31].

Важливою перевагою LSTM-архітектури над іншими типами нейронних мереж є здатність обробляти послідовності довільної довжини без необхідності паддінгу чи обрізання, що дозволяє уникнути спотворення

кореляційної структури вхідних даних. Використання двох LSTM-шарів дає змогу виявляти як короткочасні, так і довгострокові залежності в траєкторії, а також забезпечує стабільність навчання. Для запобігання перенавчанню у всіх мережах застосовується рекурентний dropout із коефіцієнтом 20% [31].

RANDI навчається на окремих мережах для різних довжин траєкторій, оскільки ефективні стратегії вилучення інформації суттєво залежать від кількості доступних точок. Для задачі класифікації навчаються шість мереж на траєкторіях фіксованих довжин $L \in \{25, 65, 125, 225, 425, 825\}$; для задачі оцінювання – чотирнадцять мереж. При обробці траєкторії довільної довжини прогноз формується як зважена комбінація двох мереж, навчених на найближчих довжинах. Кожна мережа навчається на 3×10^6 синтетичних траєкторіях із рівномірно розподіленими значеннями показника аномальної дифузії та трьома рівнями вимірювального шуму ($\text{SNR} = 10, 2$ і 1) [31].

Ключовим недоліком RANDI у контексті задачі оцінювання показника Херста є його орієнтованість на гетерогенні дані: метод навчається одночасно на п'яти моделях аномальної дифузії (CTRW, ATTM, FBM, LW, SBM) і трьох рівнях шуму. Така універсальність, необхідна для AnDi Challenge, знижує точність на однорідних траєкторіях фрактального броунівського руху [31].

2.5 Гібридний метод CINNAMON

CINNAMON (Change poInt detectionN aNd pArameter estiMation fOr aNomalous diffusion) – це гібридний метод для задачі відстеження окремих частинок, запропонований авторами [32] у рамках другого змагання AnDi Challenge. Метод вирішує три взаємопов'язані підзадачі: класифікацію типу дифузії траєкторії, виявлення точок зміни режиму руху та оцінювання параметрів дифузії в кожному сегменті між точками зміни. Принциповою відмінністю CINNAMON є поєднання трьох різнорідних підходів: глибоких

нейронних мереж, градієнтного бустингу на статистичних ознаках та аналітичних оцінювачів.

Після сегментації траєкторії показник аномальної дифузії α оцінюється в кожному сегменті окремо за допомогою градієнтного бустингу. Набір ознак для регресора є рукотворним і включає: оцінку α за допомогою TA-MSD, статистики p -варіації (для $p = 2$ та $p = 3$), середнє максимальне відхилення, емпіричну функцію автокореляції швидкостей та інші характеристики траєкторії. Саме ці ознаки – а не автоматично вивчені представлення – є основою для оцінки α , що забезпечує певний рівень інтерпретованості методу [32].

Серед переваг CINNAMON – відносна інтерпретованість завдяки аналітичним та ознаковим компонентам. Метод навчається на синтетичних даних fBm з різними фізичними моделями взаємодії і здатен обробляти траєкторії з довільною кількістю точок [32].

2.6 Порівняльний аналіз та обґрунтування потреби в новому підході

Проведений аналіз п'яти методів оцінювання показника Херста виявив ряд спільних обмежень, які створюють потребу в розробці нових підходів, здатних подолати ці недоліки. Узагальнення виявлених проблем дозволяє сформулювати вимоги до оптимального методу оцінювання та обґрунтувати доцільність розробки спеціалізованої архітектури згорткової нейронної мережі.

Найбільш критичним обмеженням усіх розглянутих методів є чутливість до довжини часового ряду. Оцінювач Уіттла та метод GHE базуються на асимптотичних властивостях, які повноцінно проявляються лише при достатньо довгих рядах. Для типових траєкторій відстеження окремих частинок, що складаються з 20-50 точок, ці методи демонструють неприйнятно велику дисперсію оцінок. Методи на основі машинного навчання RANDI та CINNAMON частково вирішують цю проблему завдяки

навчанню на синтетичних даних, однак їхня точність також суттєво знижується на коротких траєкторіях [62].

Фіксованість архітектури щодо довжини входу є ще одним спільним недоліком нейромережових підходів. Типові реалізації RANDI вимагають вхідних даних фіксованої довжини, що призводить до потреби обрізання, що спотворює статистичні властивості. Ідеальний метод повинен нативно підтримувати траєкторії змінної довжини без попередньої обробки [63].

Чутливість до адитивного шуму вимірювання є проблемою для всіх статистичних методів. Шум спотворює кореляційну структуру процесу на малих масштабах, призводячи до систематичних зміщень оцінок. Оцінювач Уіттла, GHE та метод Колмогорова-Смірнова не мають вбудованих механізмів корекції впливу шуму і вимагають додаткової попередньої обробки даних або калібрування [53].

Спеціалізація нейромережових методів на даних AnDi Challenge з п'яти різними процесами аномальної дифузії призводить до погіршення їх продуктивності на чистих fVm-траєкторіях [64]. Як показали порівняльні дослідження, RANDI демонструє найкращі результати саме на складних даних з різними типами процесів, але поступається простішим методам на однорідних наборах даних fVm. Причину цього вказали автори [65], продемонструвавши, що фрактальний рух Леві характеризується наявністю важких хвостів, і звичайні підходи не застосовні до нього на коротких траєкторіях. Автори пропонують застосування методу дробних моментів для оцінки показника Херста часових рядів фрактального руху Леві, що показує наскільки ускладнюється процес оцінювання на інших моделях дифузії [66]. Для біофізичних застосувань, де fVm є основною моделлю, такі ускладнення є суттєвим недоліком [20].

Відсутність оцінки невизначеності результату є спільною проблемою більшості розглянутих методів. Лише оцінювач Уіттла теоретично забезпечує можливість обчислення стандартних похибок, однак ці оцінки базуються на асимптотичній теорії і можуть бути ненадійними для коротких

рядів [67]. Нейромережеві методи типово видають точкові прогнози без кількісної характеристики їх надійності, що ускладнює інтерпретацію результатів у наукових дослідженнях [68].

На основі виявлених обмежень можна сформулювати ключові вимоги до оптимального методу оцінювання показника Херста для біофізичних застосувань. Гнучкість щодо довжини входу означає, що метод повинен нативно обробляти траєкторії змінної довжини без падінгу або обрізання, зберігаючи точність для коротких послідовностей від 20 точок. Робастність до шуму вимагає вбудованих механізмів корекції впливу адитивного шуму вимірювання або здатності навчитися ігнорувати шумові компоненти [69]. Спеціалізація на fVn передбачає оптимізацію методу саме для фрактального броунівського руху заради вищої точності, що втрачається при побудові універсальних методів.

2.7 Оцінювання показника Херста на основі згорткових нейронних мереж

Історично згорткові нейронні мережі були розроблені для вирішення завдань комп'ютерного зору, де вони досягли безпрецедентних успіхів завдяки двовимірним згорткам, що вилучають просторові ознаки з піксельних матриць. Проте дослідники швидко усвідомили, що фундаментальні принципи CNN – локальна зв'язність, спільне використання ваг (weight sharing) та ієрархічне формування абстракцій – ідеально підходять для обробки послідовних даних. Це призвело до стрімкого розвитку одновимірних згорткових мереж (1D-CNN), які сьогодні є загальноприйнятим підходом у класифікації та регресійному аналізі часових рядів, у тому числі для оцінювання показника Херста [70].

У структурі 1D-CNN набір згорткових фільтрів ковзає вздовж єдиної часової осі вхідного сигналу. Кожен фільтр навчається розпізнавати специфічні темпоральні патерни: від простих градієнтів та високочастотних

мікроструктурних шумів на перших шарах до складних низькочастотних макро-трендів на глибших рівнях архітектури. Оскільки показник Херста математично визначає ступінь шорсткості (або антиперсистентності) на локальному рівні та масштабування дисперсії на глобальному, згорткові фільтри природним чином резонують із механікою фрактального аналізу, фактично виконуючи роль адаптивних вейвлет-трансформацій, які налаштовуються в процесі зворотного поширення помилки [71].

Незважаючи на потужність 1D-CNN у вилученні ознак, класична архітектура стикається з критичним структурним вузьким місцем. Традиційно, після проходження сигналу через серію згорткових та пулінгових шарів (зазвичай Max Pooling для локального зменшення розмірності), формується багатоканальна карта ознак (feature map) розмірності $T' \times C$, де T' – це скорочена довжина часового ряду, а C – кількість каналів (фільтрів) останнього шару. Для того, щоб перетворити цей тензор на фінальний прогноз (наприклад, скалярну оцінку показника Херста), класичний дизайн [72] вимагає застосування шару вирівнювання (Flatten layer). Цей шар просто розгортає двовимірну матрицю в одновимірний вектор розмірності $1 \times (T' \cdot C)$ [73].

Проблема полягає в тому, що наступний за ним повнозв'язний шар багат шарового перцептрону є матрицею фіксованих ваг, яка вимагає строго визначеної кількості вхідних нейронів. Якщо довжина початкового часового ряду змінюється, змінюється й параметр T' , що робить неможливим множення на фіксовану матрицю ваг повнозв'язного шару. У реальних умовах збір даних рідко генерує ідеально однакові за довжиною послідовності. Наприклад, тривалість електроенцефалографічних записів варіюється від пацієнта до пацієнта, а потоки мережевого трафіку аналізуються динамічними вікнами. Щоб обійти цю проблему, дослідникам раніше доводилося жорстко обрізати дані (втрачаючи цінну інформацію), що спотворювало оцінку фрактальної структури. Використання рекурентних нейромереж (RNN, LSTM, GRU) частково вирішувало

проблему змінної довжини, оскільки вони обробляють дані послідовно крок за кроком. Однак LSTM-архітектури є вкрай повільними в навчанні через неможливість паралелізації обчислень і схильність надмірно фокусуватися на останніх елементах ряду, ігноруючи ранню довгострокову пам'ять [74].

Математично досконалим розв'язанням проблеми змінної довжини стало впровадження шару глобального усередненого пулінгу (Global Average Pooling, GAP), що забезпечує перехід від згорткового екстрактора ознак до багат шарового перцептрона. Спочатку запропонований у сфері комп'ютерного зору (зокрема, в архітектурах ResNet, GoogLeNet та Xception) як засіб для зменшення кількості параметрів та запобігання перенавчанню, GAP набув абсолютно нового стратегічного значення в аналізі часових рядів [70].

Механізм роботи шару GAP кардинально відрізняється від традиційного Flatten або локального Max Pooling. Математично, операція Global Average Pooling приймає на вхід карту ознак F , яка є виходом останнього згорткового шару розмірності $T \times C$. Для кожного індивідуального каналу $c \in \{1, 2, \dots, C\}$ шар обчислює середнє арифметичне значення всіх активацій вздовж усієї часової осі. Якщо позначити активацію нейрона в момент часу t для каналу c як $F_{t,c}$, то вихідне значення v_c для цього каналу формулюється рівнянням:

$$v_c = \frac{1}{T} \sum_{t=1}^T F_{t,c}, \quad (2.4)$$

де T – фактична поточна довжина карти ознак, яка прямо пропорційна довжині вхідного часового ряду [75], [76].

Результатом цієї операції є компактний вектор $v = \{v_1, v_2, \dots, v_C\}$ фіксованої розмірності $1 \times C$. Найважливішим аспектом є те, що розмірність цього вектора залежить виключно від кількості фільтрів C в останньому згортковому шарі, яка задається як гіперпараметр архітектури, і жодним чином не залежить від часової довжини T . Саме цей вектор фіксованого

розміру передається як вхідний сигнал до багатошарового перцептрона, який генерує остаточну скалярну оцінку. Даний підхід було використано для оцінки показника Херста в [70].

Архітектура згорткової нейронної мережі з використанням шару Global Average Pooling природно задовольняє вимогу гнучкості щодо довжини входу. На відміну від повнозв'язних шарів, які потребують входу фіксованого розміру, Global Average Pooling агрегує ознаки, витягнуті згортковими шарами, до вектора фіксованої розмірності незалежно від довжини вхідної послідовності. Це дозволяє обробляти траєкторії довільної довжини без модифікації архітектури [77].

Ця архітектурна парадигма має надзвичайно глибокий теоретичний резонанс із завданням оцінки показника Херста. Показник Херста за своєю природою є характеристикою, що виражає макроскопічну сталість масштабування дисперсії вздовж усього процесу. Згорткові фільтри CNN виступають у ролі детекторів специфічних мікро-патернів шорсткості або гладкості на локальних часових вікнах. Якщо аналізований процес є фрактально стабільним (наприклад, стаціонарним у приростах, як фрактальний броунівський рух), то очікувана щільність або інтенсивність прояву цих специфічних патернів буде залишатися константною незалежно від того, розглядаємо ми сегмент довжиною 500 чи 5000 точок. Відповідно, математичне усереднення цих локальних активацій за допомогою GAP ідеально агрегує інформацію, створюючи глобальний дескриптор, стійкий до часових зсувів (shift-invariant) [71].

Окрім вирішення проблеми розмірності, GAP виступає як потужний структурний регуляризатор. Біофізичні, медичні та інженерні часові ряди характеризуються високим рівнем шуму. Класичні Dense-шари, маючи мільйони параметрів, схильні до перенавчання, «визубрюючи» стохастичний шум тренувальних даних замість вивчення фундаментальних закономірностей. Оскільки сам шар Global Average Pooling не містить жодних параметрів, що навчаються, він змушує попередні згорткові шари

формувати максимально семантично наповнені та робастні карти ознак, відфільтровуючи локальні шумові сплески і залишаючи лише чистий сигнал [75], [78].

Навчання на синтетичних даних fBm з контрольованим рівнем шуму дозволяє моделі вивчити інваріантні до шуму представлення, що забезпечує робастність оцінок на зашумлених експериментальних траєкторіях [79].

Таким чином, розробка спеціалізованої CNN-архітектури для оцінювання показника Херста є обґрунтованим напрямком досліджень, який дозволяє подолати ключові обмеження існуючих методів. Поєднання гнучкості глибокого навчання з оптимізацією під задачу оцінювання на fBm-траєкторіях створює передумови для покращення точності та практичної застосовності в біофізичних дослідженнях.

Огляд наукової літератури останніх років показує, що дослідники вже зацікавилися цілеспрямованим застосуванням архітектури «1D-CNN – GAP – Регресія» для оцінки показника Херста. У своїй праці під назвою «Parameter Estimation of Long Memory Stochastic Processes with Deep Neural Networks», опублікованій у рамках престижної європейської конференції зі штучного інтелекту ECAI 2024 [70], автори представили суто неймережевий підхід до безпосередньої регресійної оцінки параметрів довгострокової пам'яті. Дослідження було зосереджене на ідентифікації показника Херста для процесів fBm та фрактального процесу Орнштейна-Уленбека, а також параметра інтеграції для моделей ARFIMA.

В основі їхньої розробки лежить модульна неймережева архітектура, спеціально розроблена для регресії секвенційних даних. Автори розробили кілька варіантів блоку трансформації послідовностей, серед яких домінуючим є саме згорткова нейронна мережа 1D-CNN. Відповідно до тексту статті, критичною особливістю їхньої архітектури є побудова так званої регресійної головки (regression head). Дослідники чітко зазначають, що регресійна головка спочатку приймає трансформовану послідовність і застосовує глобальний усереднений пулінг, що призводить

до утворення усередненого вектора ознак, який має фіксовану розмірність (таку ж, як розмір вихідної ознаки під час трансформації послідовності). Потім головка отримує кінцевий скалярний висновок з усередненого вектора ознак, використовуючи багатошаровий перцептрон.

Ця архітектурна парадигма дозволила авторам досягти видатних результатів. Для навчання моделі вони використали масштабні набори синтетичних даних високої точності. Щоб забезпечити ідеальне покриття всього простору значень показника Херста від 0 до 1, дані генерувалися за допомогою математично строгих алгоритмів: методу Девіса-Харта для fBm та методу Хоскінга для ARFIMA.

Під час тестування нейромережева модель перевершила класичні статистичні методи (зокрема метод максимальної правдоподібності Уїттла та R/S-аналіз) практично за всіма метриками: точністю, швидкістю обчислень та робастністю. Більше того, завдяки використанню GPR, автори констатують абсолютну масштабованість свого рішення. Модель, натренована на часових рядах певної довжини, продемонструвала стабільність і здатність до узагальнення під час тестування на часових рядах будь-яких інших довжин, що були відсутні в навчальній вибірці. Це підтверджує, що мережа навчилася вилучати фундаментальні фрактальні інваріанти, а не просто запам'ятала специфічні статистичні артефакти тренувального набору. У своїй роботі автори також довели, що запропонована архітектура забезпечує інваріантність до зсуву, дрейфу та масштабу.

Проте ці переваги продемонстровано на часових рядах довжиною від 10^2 до 10^4 , що зовсім не відповідає застосуванню в аналізі біознімків, де доводиться аналізувати траєкторії на порядок коротші, від 10^1 . До того ж у діапазоні довжин 200–300 запропоновані варіанти нейромереж демонстрували ефективність приблизно рівну методу Уїттла. Таким чином, існує потреба в подальшому дослідженні даного методу та його удосконаленні особливо для коротких часових рядів.

Прикладом адаптації подібних систем у клінічній медицині є дослідження Хуана Лопеса та колег, опубліковане в 2025 році [80]. У цій роботі вирішувалося завдання ранньої діагностики застійної серцевої недостатності шляхом аналізу коротких фрагментів варіабельності серцевого ритму. Фізіологічна основа методу полягає в тому, що вегетативна регуляція здорового серця генерує складні, адаптивні флуктуації, які відповідають значенням показника Херста H близько 0.5-0.7. З розвитком серцевої патології та переходом хвороби у важчі стадії, кардіоритм втрачає свою фрактальну складність, стає більш жорстким, і розподіл показника Херста зсувається в бік значень $H < 0.5$.

Для автоматизації цього аналізу дослідники розробили гібридну архітектуру, яка поєднує багат шарову CNN як екстрактор ознак та машину екстремального навчання (Extreme Learning Machine, ELM) як швидкий класифікатор. Особливістю методології було те, що мережа навчалася виключно на синтетичних монофрактальних часових рядах різної довжини (від 128 до 512), згенерованих із заданими значеннями показника Херста (від 0.1 до 0.9). При розробці архітектури пулінгу для агрегації ознак, автори [80] проаналізували традиційні перетворення для згорткових мереж, прямо перелічивши Max Pooling, Global Average Pooling, Mixed Pooling, L_p Pooling та Stochastic Pooling. Завдяки такій архітектурній композиції, модель навчилася розпізнавати патерни персистентності та антиперсистентності на синтетичних даних, а згодом була успішно застосована для інференсу на реальних фізіологічних записах RR-інтервалів (часових проміжків між послідовними скороченнями серця, що формують динамічну структуру серцевого ритму) пацієнтів. Гнучкість моделі дозволила обробляти короткі клінічні записи.

Схожий підхід до обробки біомедичних сигналів змінного розміру демонструється в дослідженні Лочани Мендіса та співавторів [81], опублікованому в журналі Scientific Reports у 2024 році. Дослідники розробили архітектуру FHR-LINet (Fetal Heart Rate – Length Invariant

Network), призначену для прогресивного оцінювання частоти серцевих скорочень плода під час пологів з метою раннього виявлення дистресу або гіпоксії. Мережа складається з паралельних та послідовних згорткових шарів, після яких вихідний тензор проходить через шар Global Average Pooling перед подачею на повнозв'язні класифікаційні шари. Автори наголошують, що саме ця архітектурна деталь робить модель інваріантною до довжини входу, дозволяючи робити передбачення на фрагментах записів кардіотокографії будь-якої варіабельної тривалості без необхідності перенавчання моделі. У цьому дослідженні показник Херста відіграв критичну роль як еталонна клінічна метрика варіабельності серцевого ритму.

Використання шару Global Average Pooling у задачах оцінки фрактальних параметрів має глибоке теоретичне та математичне обґрунтування: оскільки показник Херста є макроскопічною характеристикою самоподібності всієї системи, глобальне математичне усереднення локальних мікро-патернів шорсткості (зафіксованих згортковими фільтрами) ідеально відображає стаціонарність приростів фрактальних процесів.

3 РЕАЛІЗАЦІЯ ОЦІНЮВАЧА ПОКАЗНИКА ХЕРСТА НА ОСНОВІ CNN ТА ЕКСПЕРИМЕНТАЛЬНІ ДОСЛІДЖЕННЯ

Важливим компонентом успіху архітектури CNN + GAP для оцінки показника Херста є правильна методологія підготовки даних та налаштування метрик оцінювання. Успіх моделей глибокого навчання цілком залежить від наявності репрезентативних, збалансованих та розмічених наборів даних. Оскільки для реальних часових рядів істинне значення H завжди залишається гіпотетичним і обтяженим статистичним шумом, навчання здійснюється виключно на синтетичних траєкторіях [70].

Для створення надійних симуляцій фрактального броунівського руху доцільно скористатися алгоритмом Девіса-Харта, який генерує точні вибірки процесу зі складністю $O(N \log N)$ шляхом використання швидкого перетворення Фур'є для обчислення циркулянтних матриць. Цей генератор забезпечує створення мільйонів часових рядів зі строго заданими ідеально рівномірними значеннями $H \in (0, 1)$, що унеможлиблює проблему дисбалансу класів та зміщення нейромережі в бік найчастіших значень [82].

Експериментальні дослідження вирішено проводити виключно для коротких часових рядів довжиною від 10 до 200 часових кроків, що є актуальним при аналізі біознімків (відповідає умовам бенчмарк-датасету з аномальної дифузії AnDi-2). Дослідження складаються з трьох частин:

- у першій частині відтворюються умови експериментів із [70], досліджується запропонована ними модель з архітектурою CNN + GAP для екстремально коротких часових рядів, та досліджується ефективність запропонованих у даній кваліфікаційній роботі удосконалень даного методу;
- у другій частині досліджується стійкість методу із запропонованими удосконаленнями до шуму в даних;
- у третій частині досліджується ефективність покращень на бенчмарк-датасеті з аномальної дифузії AnDi-2.

3.1 Реалізація моделі зі згортковою мережею та шаром глобального усередненого пулінгу

Спочатку відтворено умови експерименту з [70]. Для тестування оцінювачів сформовано бенчмарк-датасет траєкторій фрактального гауссового шуму. Сітку істинних значень показника Херста взято рівномірною з кроком 0.01 у діапазоні $H \in \{0.01, 0.02, \dots, 0.99\}$, тобто 99 значень; на кожне значення H для кожної довжини $T \in \{10, 20, 30, 50, 100, 200\}$ згенеровано 300 незалежних траєкторій fGn з фіксованим seed для відтворюваності. Загалом тестовий набір містить $99 \times 300 \times 6 = 178\,200$ траєкторій. Генерація виконана через клас `FractionalGaussianNoise` бібліотеки `stochastic`, який реалізує точний алгоритм Девіса-Харта на основі циркулянтної матриці коваріацій та швидкого перетворення Фур'є.

У ролі досліджуваного нейромережевого оцінювача взято `Conv1D (n=100)` – попередньо натреновану модель з відкритого репозиторію `LMSParEst` [83], що супроводжує статтю [70] (файл вагів: «`fBm_Hurst_conv1D_n-100.pt`»). Модель навчалася авторами [70] виключно на траєкторіях довжини 100 без подальшого донавчання на коротших послідовностях, а її повністю згорткова архітектура з глобальним середнім пулінгом на виході технічно дозволяє приймати вхід довільної довжини. Архітектура моделі складається з трьох послідовних блоків. Перший – шар стандартизації, який бере різниці послідовності та нормує їх за вибічковим середнім і середньоквадратичним відхиленням; це одночасно забезпечує інваріантність до зсуву, дрейфу та масштабу, тому на вхід моделі подається саме шлях fBm (кумулятивна сума fGn), а не сам шум. Другий блок – шість згорткових шарів `Conv1D` з ядром 4, кроком 1, без доповнення; кількість каналів зростає за схемою $1 \rightarrow 64 \rightarrow 64 \rightarrow 128 \rightarrow 128 \rightarrow 128 \rightarrow 128$, після кожного шару активація `PReLU`. Третій блок – глобальне середнє пулінгу за часовою віссю та регресійна голова з двох повнозв'язних шарів (64 та 1 нейрон відповідно з активацією `PReLU`). Мінімально допустима довжина

входу цієї моделі становить 20 точок: шар стандартизації за приростами скорочує послідовність на 1 елемент, а шість послідовних згорткових шарів з ядром довжиною 4 і нульовим доповненням – ще на $6 \times 3 = 18$ елементів, тож при $T < 20$ згорткова частина не може сформувати жодного виходу. Для траєкторій довжини $T = 10$ це означає, що модель технічно непридатна до роботи; у відповідних експериментах таким траєкторіям присвоюється константна оцінка $\hat{H} = 0.5$, що дозволяє включити цю довжину у зведене порівняння на однакових тестових даних.

На відміну від [70], де порівняння здійснювалося з історичними методами (R/S-аналіз, варіограма, метод Хігучі), у даній роботі взято інший набір базових оцінювачів – таких, що широко вживаються в сучасній літературі з аналізу аномальної дифузії, та одного класичного методу для забезпечення подібності умов з результатами [70]:

- метод Уіттла (whittlehurst) – спектральний оцінювач максимальної правдоподібності, який в експериментах [70] показав найкращу якість серед класичних методів і слугує природним статистичним бенчмарком для довгих рядів;

- метод Колмогорова-Смірнова;

- CINNAMON – нейромережево-статистичний оцінювач показника аномальної дифузії α ; для випадку чистого fBm перехід до показника Херста виконується за співвідношенням $H = \frac{\alpha}{2}$;

- ruHurstEE – реалізація узагальненого показника Херста з порядком моменту $q = 2$, що відповідає класичному часово-усередненому MSD-аналізу.

Метрика якості – MSE оцінок $\hat{H}$ відносно істинного H , усереднена по всіх 29700 траєкторіях кожної довжини T . Саме така метрика була прийнята в експериментах [70], що відтворюються. Зведені результати наведено на рисунку 3.1, де відповідні криві в логарифмічних шкалах за обома осями показують залежність метрики MSE від довжини ряду T для кожного методу.

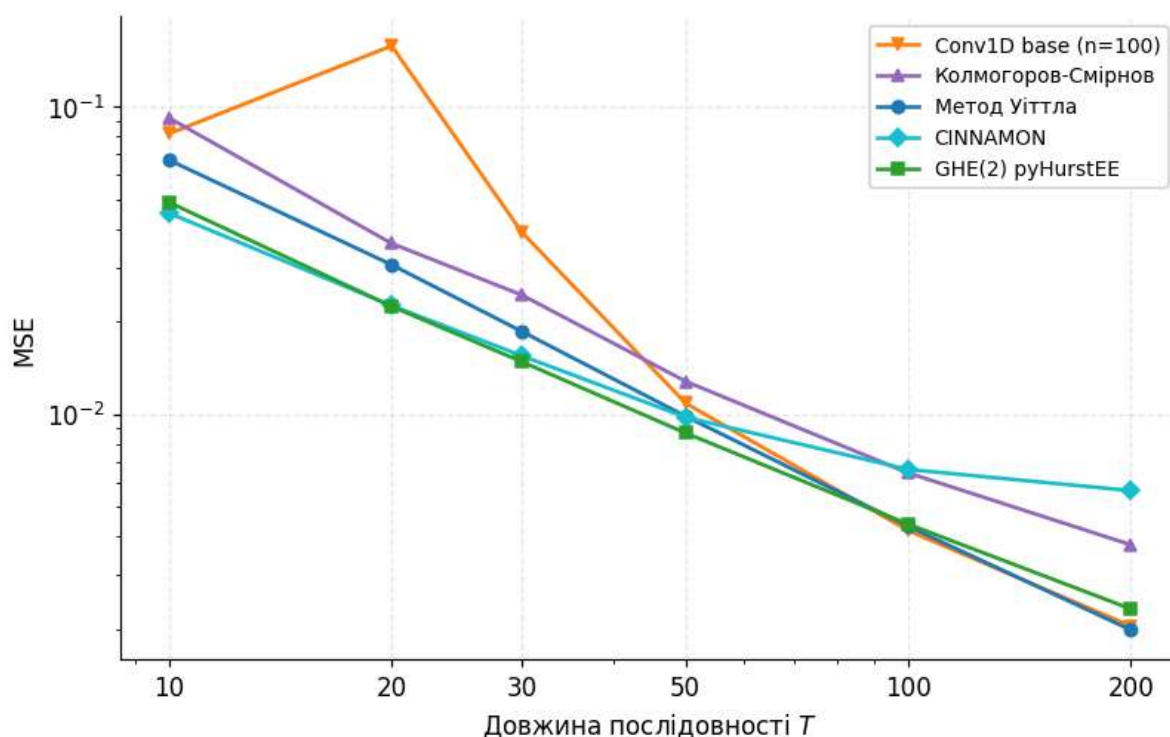


Рисунок 3.1 – Порівняння ефективності базової моделі Conv1D натренованої для довжини 100 на екстремально коротких часових рядах

Перш за все слід відзначити, що для $T = 100$ та $T = 200$ отримані значення MSE для Conv1D base та методу Уїттла практично збігаються між собою (4.179 і 4.313 при $T = 100$; 2.045 і 1.987 при $T = 200$) і добре узгоджуються зі значеннями, наведеними в таблиці 1 статті [70] (M_{conv} : 4.27 при $n = 100$ та 1.99 при $n = 200$; Уїттл: 4.33 та 2.00 відповідно). Розбіжності становлять не більше 3% і пояснюються винятково різницею в обсязі тестових вибірок та реалізаціях генератора. Це підтверджує коректність відтворення експериментального налаштування й дає підстави довіряти отриманим результатам у не дослідженому в [70] діапазоні екстремально коротких послідовностей.

У цільовому для даної роботи режимі ($T \leq 50$) крива MSE для базової моделі Conv1D демонструє чітко немонотонну поведінку, у якій варто розрізняти три якісно різні режими.

Перший – режим $T = 10$, у якому модель не приймає вхід через архітектурні обмеження і отримує тривіальну константну оцінку $\hat{H} = 0.5$. Виміряне значення $MSE = 81.667 \times 10^{-3}$ співпадає з точністю до третього знаку з теоретичною дисперсією рівномірного розподілу H на тестовій сітці $\{0.01, 0.02, \dots, 0.99\}$, яка дорівнює $(99^2 - 1)/12 \times 0.01^2 \approx 81.6 \times 10^{-3}$. Це є базовим рівнем константного оцінювача, який видає середнє значення, і виконує тут роль контрольного орієнтиру.

Другий – режим $T = 20$, що відповідає мінімально допустимій довжині входу базової моделі Conv1D. Модель тут технічно функціонує, однак виміряне значення $MSE = 157.6 \times 10^{-3}$ майже вдвічі перевищує константну межу 81.6×10^{-3} , тобто модель у цьому режимі видає систематично гірші оцінки, ніж якби вона завжди повертала 0.5. Цей результат свідчить про те, що базова Conv1D, навчена на послідовностях довжини 100, при «виродженому» рецептивному полі з 20 точок не лише не вловлює корисного сигналу, а й продукує сильно зміщений вихід.

Третій – режим $T \geq 30$, де довжина входу починає достатньо перевищувати рецептивне поле моделі, і MSE різко падає (39.1×10^{-3} при $T = 30$ проти 157.6×10^{-3} при $T = 20$), хоча досі програє pyHurstEE та методу Уїттла аж до $T = 100$, і лише при $T = 200$ досягає точності, порівнянної з еталонним методом Уїттла та найкращої серед усіх досліджуваних оцінювачів.

Натомість криві MSE для методу Уїттла, pyHurstEE, CINNAMON та Колмогорова-Смірнова в тих самих логарифмічних шкалах поведуться монотонно спадно, близько до лінійних, що відповідає очікуваному степеневому покращенню точності зі зростанням довжини ряду. У режимі $T \in \{20, 30, 50\}$ найкращу якість стабільно показує pyHurstEE (GHE з $q = 2$), а при $T = 10$ – CINNAMON. Проте CINNAMON натренований саме для екстремально коротких часових рядів, і на довжині 100 та 200 його ефективність значно погіршується. Також метод Колмогорова-Смірнова не показує жодних переваг на коротких часових рядах і суттєво поступається

іншим методам. Тому в подальшому для спрощення візуалізацій та зменшення обчислювальної складності у всіх експериментах еталонами для порівняння запропонованих модифікацій будуть два оцінювачі: метод Уіттла та pyHurstEE .

Отже, прямий перенос архітектури Conv1D + GAP, запропонованої в [70] та натренованої на послідовностях з фіксованою довжиною 100, у режим $T \leq 50$ виявляється неефективним із двох незалежних причин: жорсткого архітектурного обмеження мінімальної допустимої довжини входу ($T \geq 20$), яке робить модель непридатною для найкоротших рядів; та вираженого розподільного зсуву між довжиною, на якій модель навчалася, і довжиною, на якій вона інферить, що при граничних значеннях T призводить до точності, гіршої за константу $H = 0.5$.

Першим кроком до подолання виявлених обмежень є перенавчання тієї самої архітектури з [70] на тестовому діапазоні довжин. Якщо базова модель Conv1D, яка бачила під час навчання лише послідовності фіксованої довжини 100, демонструє різке погіршення при перенесенні на $T < 100$, то логічно перевірити, чи вирішує проблему саме мультимасштабне навчання – без жодних архітектурних модифікацій. Якщо так, то задача зводиться до правильно налаштованого тренувального пайплайну; якщо ні, то ми отримуємо нижню оцінку якості, яку взагалі можна досягти на цій архітектурі, і подальші зусилля слід спрямовувати на її модифікацію.

Тренувальний пайплайн повністю відтворює специфікацію [70] з єдиною суттєвою зміною – на кожному кроці градієнтного спуску для всього батчу обирається випадкова довжина з $\{20, 30, 50, 100, 200\}$, після чого для кожної з 512 траєкторій незалежно генерується значення H з рівномірного випадкового розподілу від 0.001 до 0.999, а також відповідна реалізація fGn методом Девіса–Харта. Кожна траєкторія додатково масштабується на множник 10^r , де r – випадкове число від -4 до 6, що формує вибірку широкого динамічного діапазону, і змушує модель явно покладатися на нормалізаційний шар при отриманні масштабно-

інваріантних оцінок. Архітектура моделі – без змін: шар стандартизації на нормування за вибіркоvim середнім та СКВ, шість шарів Conv1D з ядром 4, без доповнення, з PReLU-активаціями, глобальне середнє пулінгу, регресійна голова з двох Linear-шарів. Оптимізатор Adam з початковою швидкістю навчання 10^{-3} та косинусним розкладом, кліпінг градієнтів за нормою 5.0, функція втрат – MSE, всього 30000 кроків. Сумарний обсяг траєкторій, побачених моделлю під час навчання, становить 15.4 млн, що приблизно відповідає масштабу даних, на якому в [70] навчався Mconv (200 epoch $\times$ 100 000 траєкторій = 20 млн на основній фазі). Довжина $T = 10$ у тренувальний пайплайн не включена, оскільки знаходиться нижче за мінімально допустиму довжину входу архітектури; для оцінювання на цій довжині модель повертає константу $\hat{N} = 0.5$ як і в попередньому експерименті. Результати оцінювання перетренованої моделі Conv1D на тому самому бенчмарк-датасеті наведено в таблиці 3.1 та на рисунку 3.2.

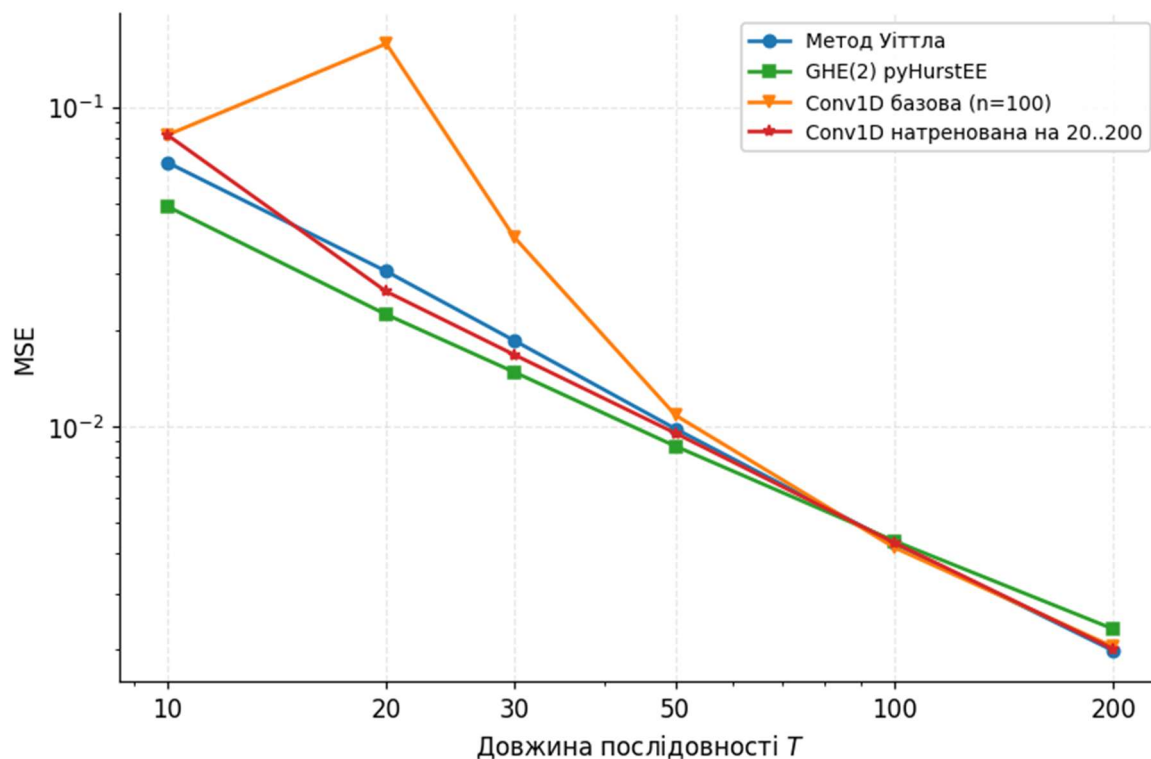


Рисунок 3.2 – Ефективність базової моделі Conv1D натренованої на прикладах від 20 до 200 часових кроків

На довжинах $T = 100$ та $T = 200$ отримана MSE становить 4.328×10^{-3} та 2.012×10^{-3} відповідно, що з точністю до 1–4% збігається як з результатом базової моделі (4.179 та 2.045×10^{-3}), так і зі значеннями з таблиці 1 статті [70] для M_{conv} (4.27 та 1.99×10^{-3}). Це підтверджує, що генератор fGn, реалізація архітектури, налаштування оптимізатора та цикл навчання працюють коректно. Тому можна зробити висновок, що експеримент, опублікований у [70], відтворений коректно.

Таблиця 3.1 – MSE оцінювачів показника Херста після перетренування архітектури з [70] на траєкторіях fGn різної довжини ($\times 10^{-3}$)

T	Метод Уіттла	pyHurstEE	Базова Conv1D з вагами з репозиторію натренована на довжині 100	Базова Conv1D перетренована на довжинах 20–200
10	66.974	48.779	81.667	81.667
20	30.58	22.404	157.631	26.435
30	18.567	14.797	39.074	16.767
50	9.831	8.673	10.84	9.526
100	4.313	4.383	4.179	4.328
200	1.987	2.325	2.045	2.012

Перетренування повністю усуває сплеск помилки при $T = 20$: MSE падає до 26.4×10^{-3} , тобто майже в шість разів. На $T = 30$ покращення також суттєве – з 39.1×10^{-3} до 16.8×10^{-3} . На $T = 50$ модель практично виходить на рівень методу Уіттла (9.53×10^{-3} проти 9.83×10^{-3}). Достатнє покриття діапазону коротких довжин у тренувальній вибірці дозволяє моделі генерувати статистично обґрунтовані оцінки. Однак перетренована модель

не випереджає кращі статистичні оцінювачі в цільовому діапазоні екстремально коротких довжин. Простий узагальнений показник Херста стабільно дає на 10–18 % меншу помилку, ніж глибока згорткова мережа зі 165 тис. параметрів, навчена на 15 мільйонах траєкторій. Тобто статистичні методи витягують з 20–50 точок суттєво більше інформації, ніж дозволяє згорткова мережа, запропонована в [70].

Графічно цей висновок виражений ще чіткіше (рисунок 3.2): крива Conv1D ($T = 10..200$) на логарифмічних шкалах злита з кривою методу Уїттла на всьому діапазоні $T \geq 20$ і йде паралельно до кривої ruHurstEE. Тобто отриманий нейромережевий оцінювач у найкращому випадку відтворює якість Уїттла.

Таким чином, експеримент з перетренуванням підтверджує коректність реалізації, але одночасно демонструє межу можливостей самої архітектури з [70]: на коротких рядах $T \leq 50$ модель з опублікованими вагами, ані перетренована її версія не переважає сучасних статистичних методів і не може бути практично рекомендована для застосувань у біофізиці, де довжини в десятки точок є типовими. Це є мотивацією ключової частини даної роботи: одного лише розширення тренувального розподілу недостатньо, і для удосконалення методу на коротких довжинах потрібно змістовно модифікувати саму архітектуру моделі.

3.2 Експеримент з адаптацією нормування до задачі аналізу часових рядів фрактального броунівського руху

Аналіз результатів попереднього підрозділу показав, що навіть при коректно налаштованому перетренуванні архітектура з [70] не випереджає кращі статистичні оцінювачі на коротких рядах. Це переводить пошук удосконалень з площини тренувального пайплайну в площину архітектурних рішень. Першим кандидатом для перегляду є шар стандартизації входу.

У архітектурі [70] нормалізація реалізована як $x = (x - \bar{x}) / \hat{\sigma}(x)$, де середнє та СКВ беруться по часовій осі окремо для кожної траєкторії. Це звичайний прийом підготовки даних для обробки нейронними мережами, може використовуватись також між шарами нейронної мережі для прискорення навчання. Однак наша гіпотеза полягає в тому, що при аналізі приростів fVn такий підхід призводить до суттєвої втрати інформації.

Гіпотеза: для траєкторій fGn вибіркове середнє приростів $\bar{x}$ несе важливу інформацію про показник Херста, особливо в режимі персистентності ($H > 0.5$) та на коротких часових рядах. Фрактальний гауссовий шум з $H > 0.5$ має повільно спадні додатні автокореляції, тому послідовність приростів схильна формувати тривалі кластери одного знаку. У такому випадку вибіркове середнє на скінченному вікні з високою ймовірністю систематично відхиляється від нуля в напрямку, що вказує на наявність кластеру одного знаку. Чим більше H , тим більша дисперсія цього відхилення відносно теоретичного нуля; при $H < 0.5$ (антиперсистентність), навпаки, прирости чергують знак, і вибіркове середнє швидше концентрується біля нуля. Отже, $\bar{x}$ емпірично корелює з H , і центрування зрізає корисну ознаку.

На рисунку 3.3 зображено запропоноване удосконалення методу оцінки показника Херста згортковою нейронною мережею з шаром глобального усередненого пулінгу через нормування без центрування. Зліва зображено оригінальний код з репозиторію [83], справа пропозиція удосконалення. Усі інші компоненти моделі залишаються без змін. Це гарантує, що будь-яка різниця в якості повністю атрибутується саме модифікації шару нормування.

Отримані в результаті криві MSE для удосконаленої моделі «Conv1D стандартизація» у порівнянні з базою «Conv1D 10..200», методом Уіттла та ruHurstEE наведено на рисунку 3.4.

Існуюче рішення в репозиторії aielte-research/LMSParEst

```
class BaseRegressor(nn.Module):
    def forward(self, x):
        if self.diff:
            x = x[:, :-1] - x[:, 1:]

        size_cond = (self.avg_base is not None)
            and (x.size(1) > self.avg_base)

        if size_cond:
            n_mult = int(x.size(1) / self.avg_base)
            x = x.view(x.size(0) * n_mult, -1)

        if self.standardize:
            m = x.mean(dim = 1, keepdim = True)
            s = x.std(dim = 1, keepdim = True)
            x = (x - m) / s
```

Запропоноване удосконалення через видалення центрування

```
class BaseRegressor(nn.Module):
    def forward(self, x):
        if self.diff:
            x = x[:, :-1] - x[:, 1:]

        size_cond = (self.avg_base is not None)
            and (x.size(1) > self.avg_base)

        if size_cond:
            n_mult = int(x.size(1) / self.avg_base)
            x = x.view(x.size(0) * n_mult, -1)

        if self.standardize:
            s = x.std(dim = 1, keepdim = True)
            x = x / s
```

Рисунок 3.3 – Запропоноване удосконалення щодо недопустимості
центрування приростів при аналізі fBm

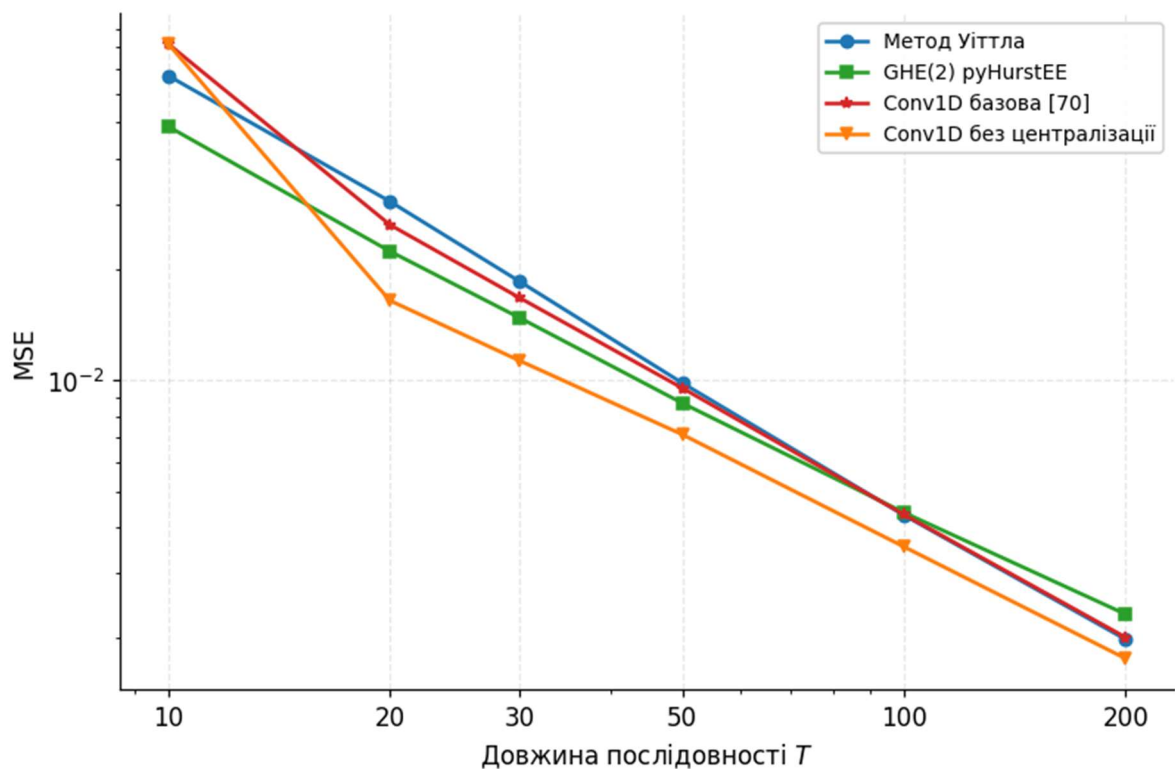


Рисунок 3.4 – Ефективність запропонованого удосконалення

Видалення центрування дає послідовне покращення на всьому діапазоні досліджуваних довжин:

- на $T = 20$ MSE падає з 26.4×10^{-3} до 16.5×10^{-3} (покращення на 38%);
- на $T = 30$ – з 16.8×10^{-3} до 11.3×10^{-3} (32%);
- на $T = 50$ – з 9.5×10^{-3} до 7.1×10^{-3} (25%);
- на $T = 100$ – з 4.3×10^{-3} до 3.5×10^{-3} (18%);
- на $T = 200$ – з 2.0×10^{-3} до 1.8×10^{-3} (12%).

Спадаючий характер відносного приросту з ростом T є самоузгодженим підтвердженням гіпотези: чим довший ряд, тим менший вклад вибіркового середнього як ознаки порівняно з основним сигналом, який модель витягує з кореляційної структури приростів, і тим менший виграш від його збереження. Натомість на коротких рядах, де цінна кожна додаткова ознака, центрування виявилось прямою втратою інформації.

Принципово важливо, що отримана модель уперше в даному дослідженні випереджає $ruHurstEE$ на всіх довжинах, де згорткова мережа технічно функціонує. На рисунку 3.4 крива нової моделі лежить нижче від обох статистичних бенчмарків на всьому діапазоні $T \geq 20$, причому розрив у логарифмічних шкалах залишається приблизно сталим – тобто покращення не зникає зі зростанням довжини.

Окремого коментаря заслуговує випадок $T = 10$, який залишається непідвладним ані базовій, ані модифікованій версії: рецептивне поле згорткової частини в 19 точок робить модель архітектурно нездатною прийняти 10-точковий вхід, незалежно від схеми нормування. Подолання цього обмеження виходить за межі модифікації шару стандартизації і потребує перегляду самої згорткової частини – цьому присвячено наступний підрозділ.

Сумарно експеримент дозволяє сформулювати такий висновок. Центрування приростів у схемі [70] є шкідливим при роботі безпосередньо з приростами fGn : воно зрізає ознаку, яка несе самостійну інформацію про показник у режимі коротких рядів. Перехід від повної стандартизації до

масштабного нормування забезпечує трансформацію нейромережевого оцінювача: замість результатів, порівнянних із методом Уіттла, модель демонструє стабільну перевагу над усіма статистичними еталонами в діапазоні $T \in [20, 200]$.

Це є першою з модифікацій, яка фактично виводить запропонований у [70] метод за межі досягнутої там якості й робить його придатним для застосування на коротких часових рядах, типових для аналізу спостережень одиноких частинок.

3.3 Експерименти з ruHurstEE як екстрактором ознак

Усі модифікації, розглянуті в підрозділах 3.1–3.2, стосувалися виключно нейромережевої складової оцінювача – шару нормування, форми регресійної голови, передачі довжини ряду. При цьому кращим статистичним бенчмарком на коротких рядах послідовно залишався ruHurstEE – реалізація узагальненого показника Херста другого порядку (GHE з $q = 2$). Це підказує наступний крок: замість того щоб змагатися зі статистичним методом, варто використати його як вбудований фізично інформований екстрактор ознак, поряд зі згортковим, та об'єднати їх вихід на рівні регресійної голови. Тобто від конкуренції між двома сімействами оцінювачів перейти до їх поєднання в єдиній архітектурі.

Ідея фізично інформованих ознак не нова в задачах оцінювання параметрів стохастичних процесів. Її логіка проста: статистичний оцінювач, побудований на конкретній параметричній моделі (тут – степеневій залежності MSD від лагу для fGn), вносить у модель інтерпретовану інформацію, яку згорткова мережа змушена була б виводити з даних з нуля. У результаті модель витрачає менше ресурсів на повторне відкриття вже відомих закономірностей, і залишається більше ресурсів на корекцію систематичних похибок самого статистичного оцінювача та на врахування ефектів короткої довжини.

До архітектури з підрозділу 3.2 (нормування без центрування, шість шарів Conv1D з ядром 4, GAP, двошарова MLP-голова) додається паралельний фізично інформований екстрактор – модуль pyHurstEE з порядком моменту $q = 2$ та лаговим вікном до 4 точок. Він обчислюється безпосередньо за вхідним рядом fGn до подачі в згорткову частину і повертає скалярну оцінку $\hat{H}$. Цей скаляр конкатенується зі 128-вимірним вектором ознак, отриманим з GAP згорткової частини, утворюючи об'єднаний 129-вимірний вектор. Перший лінійний шар MLP-голови, відповідно, має 129 входів. Згортковий екстрактор та фізично інформований екстрактор працюють паралельно й незалежно, MLP-голова навчається їх правильно поєднувати на основі градієнта від функції втрат.

Також було випробувано розширення попередньої конфігурації – до 129-вимірного об'єданого вектора додатково конкатенується скаляр $\log T$, де T – довжина даного вхідного ряду. Перший лінійний шар MLP-голови приймає 130 входів. Решта елементів архітектури й весь тренувальний пайплайн (30000 кроків, батч 512, мультимасштабна вибірка довжин з $\{20, 30, 50, 100, 200\}$) ідентичні попереднім підрозділам. Модель оперує траєкторіями змінної довжини, при цьому, як видно з результатів попередніх досліджень [70], величина зміщення суттєво залежить від T , а інформація про конкретну довжину входу губиться після глобального середнього пулінгу. Вирішено перевірити, чи допоможе моделі явне передавання довжини ряду як додаткової ознаки на вхід регресійної голови.

Криві MSE для нової конфігурації у порівнянні з кращою моделлю попередніх підрозділів (Conv1D без централізації) та статистичними бенчмарками наведено на рисунку 3.5. Додавання фізично інформованого екстрактора без передавання довжини ряду дає неоднорідну картину:

- на $T = 20$ і $T = 30$ ефективність методу є мінімальною;
- на $T = 50$ з'являється покращення з 7.13×10^{-3} до 6.806×10^{-3} (4.5%);
- на $T = 100$ – з 3.54×10^{-3} до 3.392×10^{-3} (4.2%);
- на $T = 200$ – з 1.763×10^{-3} до 1.717×10^{-3} (2.6%).

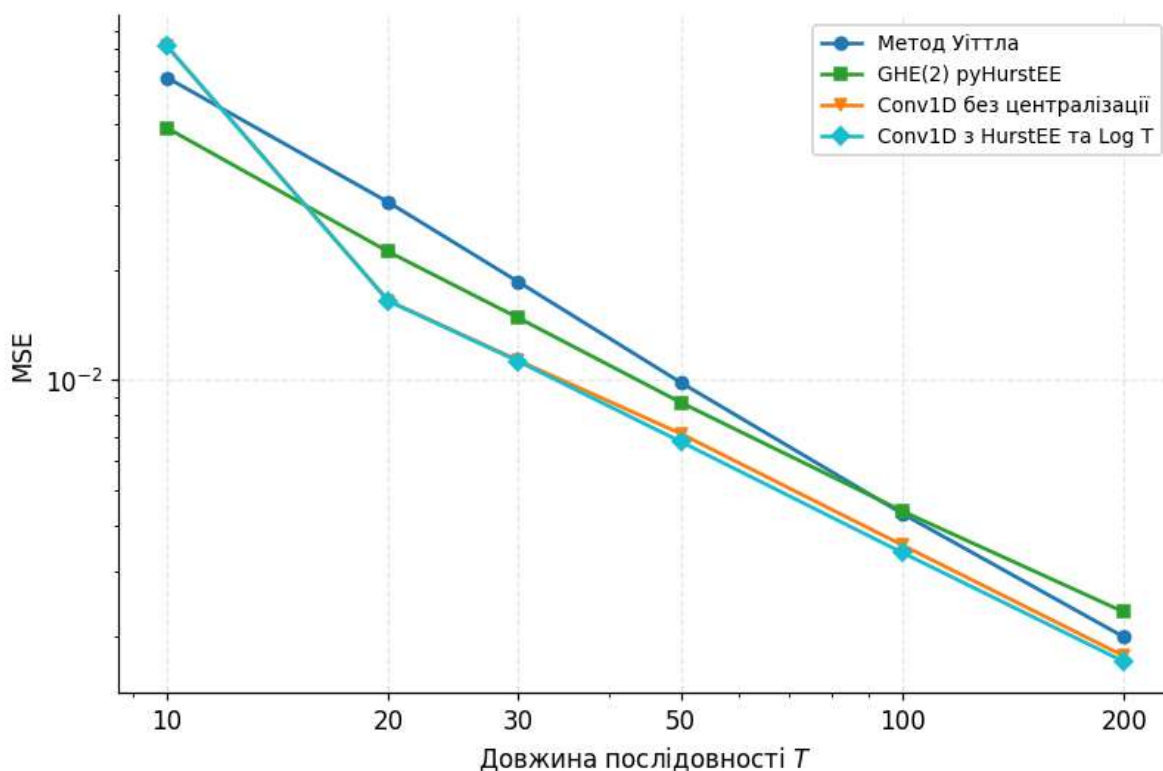


Рисунок 3.5 – Ефективність запропонованого поєднання з руHurstEE

Тобто сам руHurstEE як фізично інформований екстрактор починає допомагати з $T \geq 50$, тоді як на найкоротших довжинах (20–30 точок) його сигнал не вносить нової інформації.

Конфігурація з додатковим передаванням $\log T$ дає додаткове покращення на всіх довжинах, де модель функціонує. Виграш, порівняно з додаванням лише руHurstEE, складає 0.3–1.3%, залежно від T. Цей ефект слабко, але стабільно проявляється: явна передача довжини ряду в голову допомагає, проте лише в межах одного відсотка. Порівняно з моделлю підрозділу 3.2 кінцева конфігурація виграє на всіх довжинах.

Об'єднання згорткового та фізично інформованого екстракторів через спільну MLP-голову працює там, де самостійні статистичні методи перестають бути джерелом шуму і починають надавати моделі систематичну компоненту, не повністю передбачувану згорткою. Для

коротких рядів $T \leq 30$ згорткова частина вже витягнула з даних практично всю наявну інформацію, і `pyHurstEE` як зашумлений оцінювач для цих довжин не додає корисного сигналу. Для рядів $T \geq 50$ фізично інформована ознака стабільно покращує якість на 3–5%.

Підсумовуючи, конфігурація з додаванням ознак `pyHurstEE` та довжини приймається як основна за результатами підрозділів 3.1–3.3 і слугує основою для подальших експериментів, зокрема архітектурного зняття обмеження $T \geq 20$.

3.4 Спрощення архітектури згорткового екстрактора ознак

Для опрацювання довжин рядів від 10 часових кроків необхідно спростити архітектуру згорткової частини. Розмір ядра 4 доцільно залишити лише на першому кроці, і побудувати чотиришаровий екстрактор за схемою кількості ядер 4–3–3–3. На вхід такого екстрактора можна подати часовий ряд мінімальною довжиною 10. Компенсувати зменшення кількості шарів можна збільшуючи кількість каналів (виходів шарів) вдвічі. Також можливо додати шар в MLP.

Мінімальна довжина входу впливає з обраної схеми ядер. Згортковий екстрактор скорочує довжину на $(4 - 1) + (3 - 1) + (3 - 1) + (3 - 1) = 9$ елементів. Отже, для отримання хоча б одного елемента на виході згорткової частини достатньо вхідного ряду довжиною 10 точок, що точно відповідає найкоротшій довжині цільового діапазону. Подальше скорочення глибини мережі або зменшення першого ядра до 3 точок робить архітектурно можливим прийом ще коротших рядів, але одночасно зменшує локальне рецептивне поле першого шару нижче за межу, на якій згортка ще може вловити нетривіальні часові залежності, тому варіант 4–3–3–3 приймається як компромісний.

Зменшення глибини з шести шарів до чотирьох потенційно знижує ємність моделі. Щоб компенсувати цей ефект, кількість вихідних каналів

подвоюється порівняно з [70], а саме $1 \rightarrow 128 \rightarrow 256 \rightarrow 256 \rightarrow 256$ замість $1 \rightarrow 64 \rightarrow 64 \rightarrow 128 \rightarrow 128 \rightarrow 128 \rightarrow 128$. Розмірність вектора ознак, що подається в MLP-голову після глобального середнього пулінгу, відповідно подвоюється з 128 до 256. Регресійна голова також розширюється: до неї додається ще один лінійний шар, і вона набуває форми $\text{Linear}(258 \rightarrow 256) \rightarrow \text{PReLU} \rightarrow \text{Linear}(256 \rightarrow 64) \rightarrow \text{PReLU} \rightarrow \text{Linear}(64 \rightarrow 1)$. Перший шар приймає 258 входів – 256-вимірний вектор згорткової частини та скаляр оцінки ruHurstEE та скаляр $\log T$, що повністю зберігає схему об'єднання згорткового та фізично інформованого екстракторів із підрозділу 3.3.

Тренувальний пайплайн розширюється на нову довжину: набір довжин, з якого випадково обирається довжина чергового батчу, тепер становить $\{10, 20, 30, 50, 100, 200\}$, тобто $T = 10$ включається в навчання нарівні з рештою. Решта налаштувань – 30000 кроків, батч 512, оптимізатор Adam з косинусним розкладом, повністю відповідають попереднім підрозділам.

Криві MSE отриманої моделі Conv1D зі спрощеним екстрактором у порівнянні з попередньою найкращою конфігурацією та статистичними бенчмарками наведено на рисунку 3.6.

Модель уперше в усіх експериментах даної роботи дає змістовну оцінку для $T = 10$. MSE падає з константного значення 81.667×10^{-3} , що відповідало стратегії «завжди повертати 0.5», до 30.955×10^{-3} – у 2.6 рази менше. Це також у 1.6 рази краще за ruHurstEE , який на цій довжині давав 48.779×10^{-3} , і в 2.2 рази краще за метод Уїттла (66.974×10^{-3}). Тобто згорткова мережа з рецептивним полем, точно адаптованим до мінімальної довжини, переважає всі досліджувані статистичні оцінювачі на найкоротших рядах.

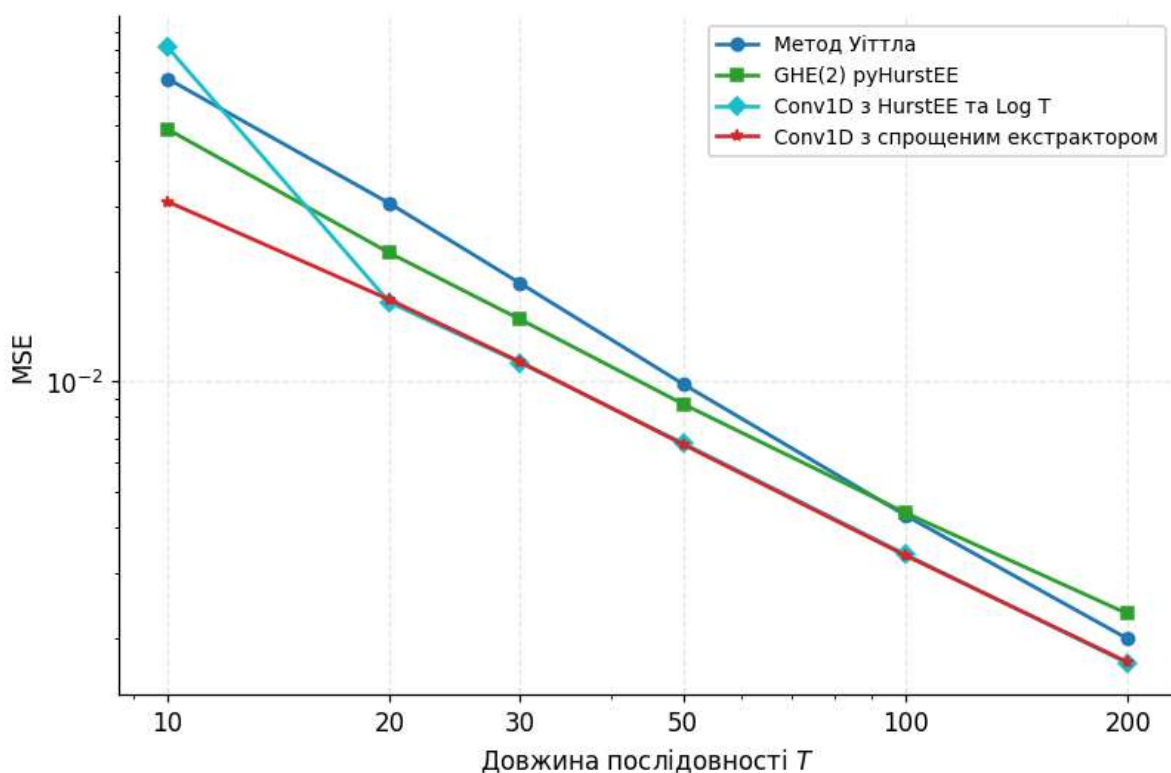


Рисунок 3.6 – Результат спрощення архітектури згорткового екстрактора

Не менш важливо, що розширення тренувального розподілу на $T = 10$ і архітектурне спрощення згорткової частини не погіршили якість оцінювача на жодній іншій довжині.

Конфігурація моделі, отримана в цьому підрозділі, приймається як підсумкова архітектура запропонованого в роботі удосконалення оцінювача [70]: шар нормування $\frac{x}{\sigma(x)}$ без центрування, чотиришаровий згортковий екстрактор з ядрами 4–3–3–3 й каналами 128–256–256–256, паралельний детермінований pyHurstEE-екстрактор, конкатенація 256-вимірного вектора згортки, скалярної оцінки pyHurstEE та $\log T$ у вхід трьохшарової MLP-голови ($258 \rightarrow 256 \rightarrow 64 \rightarrow 1$). Ця модель уперше серед розглянутих у роботі забезпечує перевагу над усіма досліджуваними статистичними оцінювачами на всьому цільовому діапазоні $T \in [10, 200]$ і слугує основою для подальших експериментів зі стійкості до шуму та застосуванням на бенчмарк-датасеті AnDi-2.

3.5 Дослідження робастності моделі до адитивного шуму

Для об'єктивної оцінки робастності запропонованої моделі до зашумлених даних проведено навчання на синтетичних даних fBm з контрольованим рівнем шуму. Результати порівняно з моделлю, навченою на чистих даних у підрозділі 3.4. Для валідації створено датасети аналогічні бенчмарк-датасету з підрозділу 3.1, що використовувався у всіх дослідях 3.1–3.4. Додано п'ять датасетів такого ж розміру та з такими ж параметрами, і єдиною зміною є рівень адитивного шуму по відношенню до середньоквадратичного відхилення основного сигналу. Цей рівень вибрано з ряду $[0.01, 0.1, 0.3, 0.5, 0.9]$. До датасету додано результати оцінки методами Уїттла, GHE(2) та CINNAMON.

До кожного прикладу навчального набору даних додається білий гауссів шум, СКВ якого рівномірно вибирається з інтервалу від 0 до заданого максимального значення $\sigma_{\max}$. Для випадків значень $\sigma_{\max} < 0.2$ не спостерігалось значних відмінностей від навчання на чистих часових рядах. Результати для $\sigma_{\max} = 0.3$ та $\sigma_{\max} = 0.6$ наведені на рисунках 3.7–3.9.

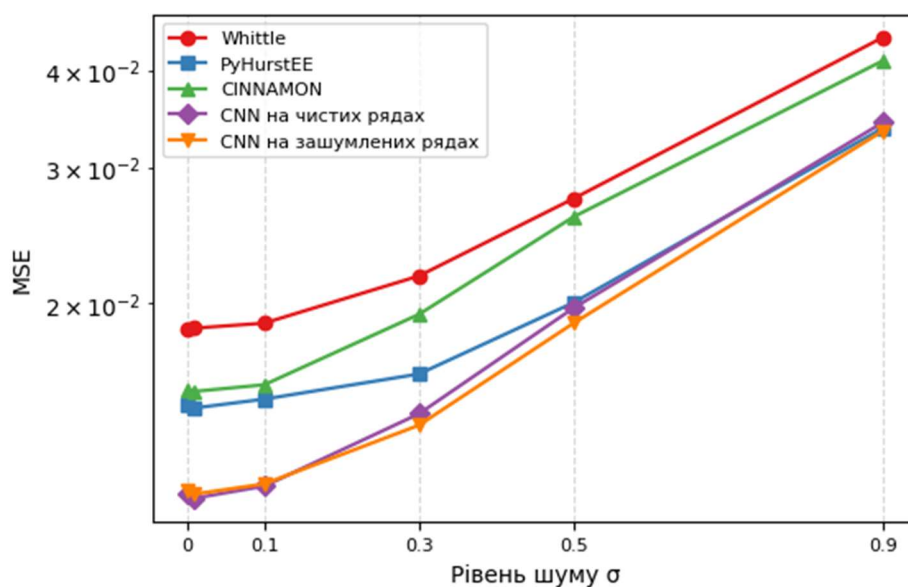


Рисунок 3.7 – Результат на рядах довжиною 30 моделі, навченої на прикладах з максимальним рівнем доданого шуму 0.3

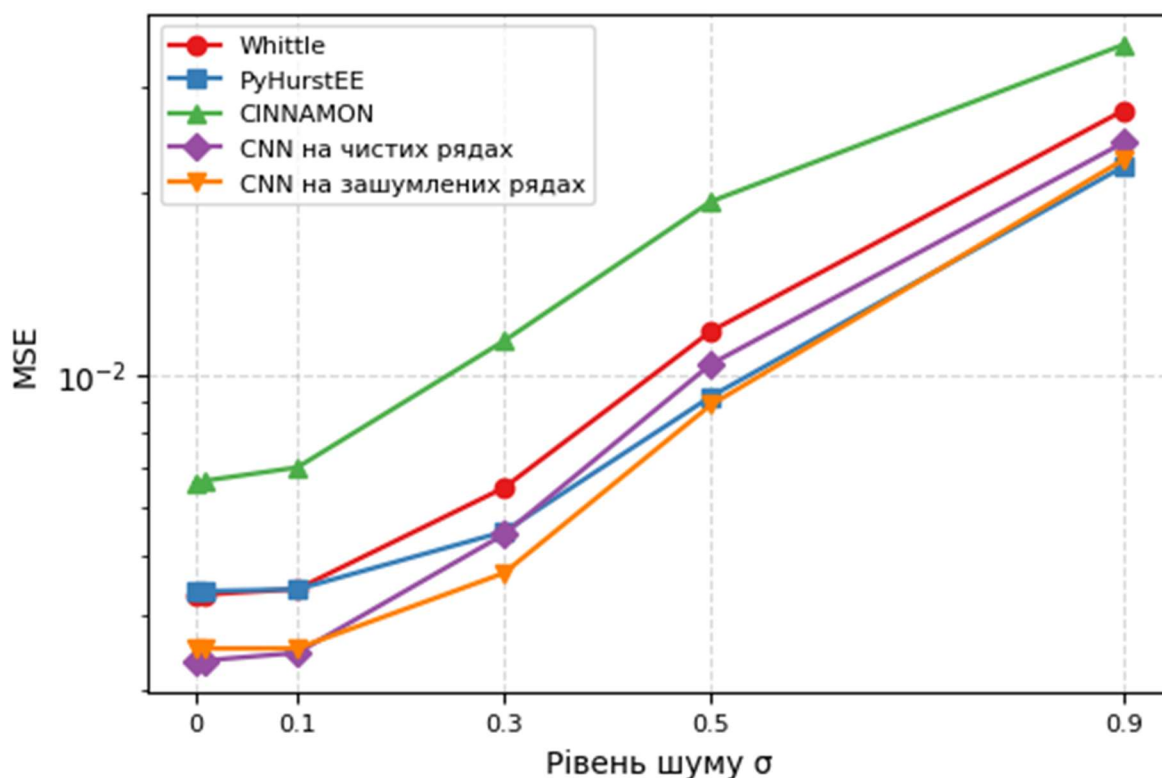


Рисунок 3.8 – Результат на рядах довжиною 100 моделі, навченої на прикладах з максимальним рівнем доданого шуму 0.3

Аналіз рисунку 3.7 показує, що на довжині рядів 30 при $\sigma_{\max} = 0.3$ криві MSE моделі, навченої на зашумлених рядах, незначно відрізняються від кривих моделі, навченої на чистих рядах. При цьому, при рівні шуму до 0.1, кращий результат у моделі CNN на чистих рядах, а при вищому шумі – у CNN на зашумлених рядах. На коротких рядах $T < 50$ модель, навчена з $\sigma_{\max} = 0.3$, демонструє робастність і на рядах з шумом 0.5, тобто вищим, ніж була навчена.

Аналіз рисунка 3.8 показує результати при збільшенні довжини ряду. На довжині рядів $T = 100$ при $\sigma_{\max} = 0.3$ криві MSE моделі, навченої на зашумлених рядах, відрізняються від кривих моделі, навченої на чистих рядах, в основному лише коли рівень шуму дорівнює 0.3, а на решті діапазонів довжин T та валідаційних рівнів шуму σ криві обох моделей близькі між собою. У даному випадку при збільшенні довжини ряду модель

не демонструє робастності на рядах з шумом 0.5 (вищому, ніж шум, на якому була навчена).

Водночас слід відзначити, що стратегія з невисоким рівнем шуму навчання $\sigma_{\max} = 0.3$ не має суттєвого погіршення точності оцінювання на чистих даних.

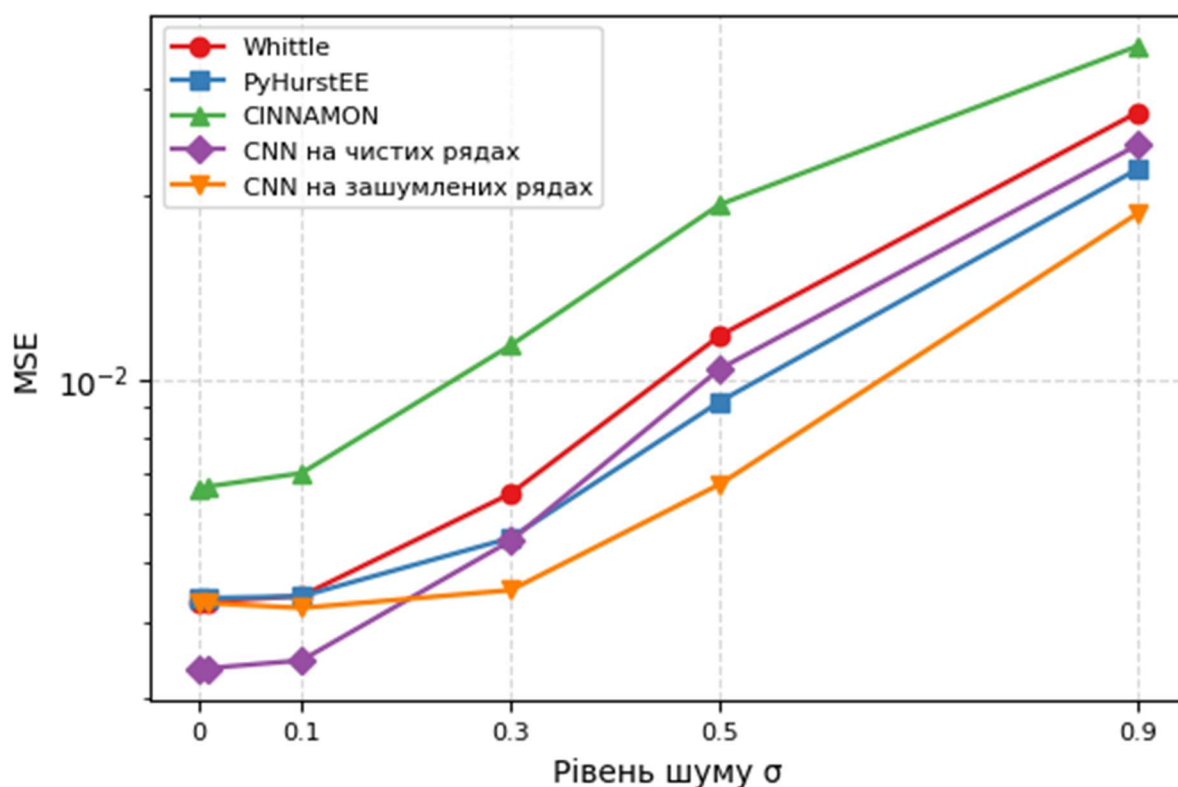


Рисунок 3.9 – Результат на рядах довжиною 100 моделі навченої на прикладах з максимальним рівнем доданого шуму 0.6

Натомість для $\sigma_{\max} = 0.6$ картина якісно змінюється (рисунок 3.9). Стає видимим компроміс між точністю оцінювача на чистих даних і його стійкістю до шуму. Метод показує свою ефективність при рівні шуму валідаційних даних вище 0.2. У той же час при рівні шуму валідаційних даних нижче 0.2 (тобто на досить чистих) модель, навчена з адитивним шумом, значно поступається не тільки моделі на чистих даних, а й іншим оцінювачам показника Херста.

Схожий компроміс між робастністю та точністю спостерігався в [79] при аналізі оцінювачів показника аномальної дифузії методом TA-MSD. При цьому аналогічний метод узагальненого показника Херста, як і метод Уїтгла, у нашому експерименті показує більш високу робастність, ніж CINNAMON та запропонований метод.

Експерименти з контрольованою шумовою аугментацією дозволяють сформулювати такий висновок: навчання на рядах з доданим шумом забезпечує робастність на деякому діапазоні шуму, погіршуючи точність на чистих даних. Робастність не виявилась сильною стороною запропонованого методу. Навіть включення екстрактора `pyHurstEE` не забезпечило робастності подібної до спостереженої у самому методі `pyHurstEE`. Крім того, навчання з доданим шумом невисокого рівня не узагальнюється на більш високі рівні шуму. Тому максимальний рівень шуму по відношенню до основного сигналу слід вибирати як очікуваний рівень шуму в даних, які досліджуються.

3.6 Порівняльний експеримент на датасеті AnDi-2

Для об'єктивної оцінки розробленого методу та порівняння його з еталонними підходами було проведено фінальний експеримент на бенчмарк-датасеті аномальної дифузії AnDi-2 [84]. Він містить синтетичні гетерогенні траєкторії фрактального броунівського руху зі змінними режимами дифузії та точками зміни, що відповідає реальним умовам відстеження окремих частинок у живих клітинах. Вибір цього датасету обумовлений основними чинниками: він є визнаним галузевим стандартом для об'єктивного порівняння методів; водночас, спеціалізований метод CINNAMON [32] був розроблений під час другого змагання AnDi Challenge, і справедливе порівняння запропонованого в роботі методу з CINNAMON можливе тільки на датасеті AnDi-2. Початковий етап аналізу даного датасету було апробовано на конференціях [85] та [86].

Метрикою порівняння обрано середню абсолютну похибку (MAE) оцінювання показника Херста на тестових сегментах датасету. MAE є стандартною метрикою в задачах оцінювання показника аномальної дифузії і використовувалася в AnDi-2 Challenge для показників аномальної дифузії, які пов'язані з показником Херста як $H = \frac{\alpha}{2}$. Крім того, MAE надає безпосередньо інтерпретовані результати в тих самих одиницях, що і сам показник H .

У порівнянні беруть участь чотири методи. Метод Уіттла є параметричним спектральним оцінювачем, детально описаним у підрозділі 2.1; він включений до порівняння як класичний статистичний базисний рівень. Метод TA-MSD з відсіканням 3% є модифікованим методом часово-усередненого середньоквадратичного зміщення з відсіканням трьох відсотків найбільших середньоквадратичних зміщень для забезпечення стійкості до викидів [79]. Він включений як другий статистичний базовий підхід. CINNAMON [32] є гібридним методом, спеціалізованим для задачі оцінювання параметрів у другому AnDi Challenge; він є цільовим конкурентом для нейромережевого підходу, розробленого в цій роботі. Метод Conv+GAP зі спрощеною архітектурою є фінальною версією нейромережі CNN+GAP, розробленою й навченою в рамках даної роботи згідно з підходом, описаним у підрозділі 3.4.

Результати порівняльного експерименту представлені в таблиці 3.2. Отримані результати демонструють чітку ієрархію між методами. Класичний оцінювач Уіттла, попри свою строгу теоретичну обґрунтованість та асимптотичну ефективність, показує найгірший результат ($MAE = 0.2014$), що підтверджує висновки підрозділу 2.1 щодо суттєвого погіршення його точності на коротких сегментах траєкторій. Метод TA-MSD з відсіканням 3% надає помірне покращення ($MAE = 0.1727$).

Таблиця 3.2 – Порівняння методів на датасеті AnDi-2 (MAE)

Метод	MAE
Метод Уіттла	0.2014
TA-MSDt з відсіканням 3% [79]	0.1727
CINNAMON	0.1648
Conv+GAP (спрощена архітектура згідно з підрозділом 3.4)	0.1633

Гібридний метод CINNAMON, спеціально розроблений для задач AnDi-2 Challenge, демонструє MAE = 0.1648, що є значним покращенням відносно класичних підходів. Розроблений у даній роботі метод Conv+GAP зі спрощеною архітектурою досягає найкращого результату серед усіх порівнюваних методів – MAE = 0.1633. Перевага над CINNAMON становить 0.0015 за абсолютним значенням. Попри те, що ця різниця є відносно невеликою, цей результат є важливим. CINNAMON є спеціалізованим методом, цілеспрямовано спроектованим під задачі AnDi із використанням галузевих знань у конструкції ознак, тоді як запропонована архітектура є загальною, натренованою на синтетичних рядах без доданого шуму, та не містить жодних закладених припущень про специфіку датасету – ознаки витягуються автоматично в процесі навчання.

Таким чином, результати порівняльного експерименту підтверджують дві центральні гіпотези роботи: нейромережева архітектура Conv+GAP перевершує класичні методи оцінювання показника Херста (Уіттла та TA-MSD) на гетерогенних траєкторіях бенчмарк-датасету AnDi-2, і досягає кращих результатів, ніж спеціалізований метод CINNAMON, що є прямим підтвердженням ефективності запропонованого підходу та ролі шару Global Average Pooling.

ВИСНОВКИ

У кваліфікаційній роботі розв'язано актуальну наукову задачу удосконалення методу оцінювання показника Херста часових рядів, симульованих фрактальним броунівським рухом, на основі згорткової нейронної мережі з шаром глобального усередненого пулінгу. Завдання на кваліфікаційну роботу виконано в повному обсязі: проаналізовано предметну галузь та існуючі статистичні й нейромережеві методи оцінювання показника Херста, спроектовано експериментальне середовище, синтезовано датасет траєкторій фрактального гауссового шуму з відомими значеннями H , реалізовано оцінювач показника Херста на основі згорткової нейронної мережі та проведено комплексний експериментальний аналіз із порівнянням розробленого підходу з аналогами на власних синтетичних даних та бенчмарк-датасеті AnDi-2.

За результатами експериментального дослідження отримано низку чисельних показників, що підтверджують ефективність запропонованих архітектурних удосконалень оцінювача:

- видалення центрування в шарі нормалізації входу забезпечує зменшення середньоквадратичної похибки оцінювання показника Херста на 12–38% порівняно з опублікованою базовою архітектурою;

- додавання паралельного фізично інформованого екстрактора ruHurstEE на основі узагальненого показника Херста другого порядку забезпечує стабільне покращення точності оцінювання в межах 3–5 % для довжин $T \geq 50$, а явна передача довжини ряду у вигляді $\log T$ дає додатковий приріст у межах 0.3–1.3 %;

- спрощення згорткового екстрактора за схемою ядер 4–3–3–3 із подвоєною кількістю каналів дозволяє отримувати змістовну оцінку показника Херста на надкоротких рядах довжиною $T = 10$ точок;

- на бенчмарк-датасеті AnDi-2 запропонована удосконалена архітектура Conv+GAP досягає середньої абсолютної похибки оцінювання

показника Херста $MAE = 0.1633$, що є найкращим результатом серед усіх досліджуваних методів.

Якісним показником досягнутих результатів є те, що запропонована в даному дослідженні архітектура рівномірно випереджає всі обрані статистичні бенчмарки на цільовому діапазоні $T \in [10, 200]$. Таким чином, мету роботи – удосконалення методу оцінювання показника Херста часових рядів будь-якої довжини, симульованих фрактальним броунівським рухом, на основі згорткової нейронної мережі з шаром глобального пулінгу – досягнуто, а всі поставлені в завданні задачі виконано в повному обсязі.

Основними новими науковими результатами роботи є:

- обґрунтовано та експериментально підтверджено принцип відмови від центрування вибіркового середнього в шарі нормалізації входу нейромережових оцінювачів показника Херста: показано, що вибіркове середнє приростів фрактального гауссового шуму корелює з показником H на скінченних вікнах і несе самостійну інформацію, втрата якої погіршує оцінювання саме на коротких рядах, де приріст ознак є критично важливим;
- розроблено гібридну архітектуру, у якій згортковий екстрактор ознак поєднано з паралельним детермінованим фізично інформованим оцінювачем $ruHurstEE$ на основі узагальненого показника Херста другого порядку, що зменшує необхідність повторного відкриття мережею вже відомих статистичних закономірностей і дозволяє стабільно покращувати точність на довжинах $T \geq 50$;
- запропоновано спрощену схему згорткового екстрактора з ядрами 4–3–3–3 та подвоєною кількістю каналів, яка дозволяє нейромережевому оцінювачу показника Херста стабільно працювати на рядах мінімальної довжини $T = 10$ точок та переважати всі досліджувані статистичні методи на цьому діапазоні.

Код та результати дослідів для можливості їх відтворення опубліковано у відкритому репозиторії [87].

Отримані в роботі результати мають прикладну цінність, оскільки суттєво покращено метод оцінювання, що опублікований у [70] та застосовується зокрема в застосунках ранньої діагностики застійної серцевої недостатності [80], прогресивного оцінювання частоти серцевих скорочень плода під час пологів з метою раннього виявлення дистресу або гіпоксії [81], виявлення епілептичних нападів [88], мультифрактального аналізу ЕЕГ-каналів для класифікації стадій сну та ідентифікації когнітивних навантажень або рухової уяви [89], детекції несправностей обладнання [78]. Упровадження запропонованого удосконалення у вигляді відсутності центрування вхідних часових рядів та додавання паралельного фізично інформованого екстрактора ruHurstEE на основі узагальненого показника Херста другого порядку значно підвищить точність результатів у вказаних застосуваннях методу.

Таким чином, у кваліфікаційній роботі повною мірою досягнуто поставленої мети, виконано всі сформульовані завдання, отримано нові науково-практичні результати, що мають теоретичну та прикладну цінність.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Anomalous diffusion models and their properties: non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking / R. Metzler et al. *Phys. Chem. Chem. Phys.* 2014. Vol. 16, no. 44. P. 24128–24164. URL: <https://doi.org/10.1039/c4cp03465a> (date of access: 26.04.2026).
2. Einstein A. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik.* 1905. Vol. 322, no. 8. P. 549–560. URL: <https://doi.org/10.1002/andp.19053220806> (date of access: 26.04.2026).
3. Krapf D. Mechanisms underlying anomalous diffusion in the plasma membrane. *Lipid Domains.* 2015. P. 167–207. URL: <https://doi.org/10.1016/bs.ctm.2015.03.002> (date of access: 26.04.2026).
4. Mandelbrot B. B., Van Ness J. W. Fractional Brownian motions, fractional noises and applications. *SIAM Review.* 1968. Vol. 10, no. 4. P. 422–437. URL: <https://doi.org/10.1137/1010093> (date of access: 26.04.2026).
5. Fuliński A. Fractional Brownian motions: memory, diffusion velocity and correlation functions. *J. Phys. A: Math. Theor.* 2017. Vol. 50. P. 054002.
6. Saxton M. J., Jacobson K. Single-particle tracking: applications to membrane dynamics. *Annual Review of Biophysics and Biomolecular Structure.* 1997. Vol. 26, no. 1. P. 373–399. URL: <https://doi.org/10.1146/annurev.biophys.26.1.373> (date of access: 26.04.2026).
7. Seier K. Investigation of dynamic processes of prototypical class A GPCRs by single-molecule microscopy : doctoral thesis. 2020. URL: <https://nbn-resolving.org/urn:nbn:de:bvb:20-opus-199739> (date of access: 26.04.2026).
8. Höfling F., Franosch T. Anomalous transport in the crowded world of biological cells. *Reports on Progress in Physics.* 2013. Vol. 76, no. 4. P. 046602. URL: <https://doi.org/10.1088/0034-4885/76/4/046602> (date of access: 26.04.2026).

9. Di Matteo T., Aste T., Dacorogna M. M. Scaling behaviors in differently developed markets. *Physica A: Statistical Mechanics and its Applications*. 2003. Vol. 324, no. 1-2. P. 183–188. URL: [https://doi.org/10.1016/s0378-4371\(02\)01996-9](https://doi.org/10.1016/s0378-4371(02)01996-9) (date of access: 26.04.2026).
10. Objective comparison of methods to decode anomalous diffusion / G. Muñoz-Gil et al. *Nature Communications*. 2021. Vol. 12, no. 1. URL: <https://doi.org/10.1038/s41467-021-26320-w> (date of access: 26.04.2026).
11. The anomalous diffusion challenge: single trajectory characterisation as a competition / G. Muñoz-Gil et al. *Emerging Topics in Artificial Intelligence 2020*, Online Only, United States, 24–28 August 2020 / ed. by G. Volpe et al. 2020. URL: <https://doi.org/10.1117/12.2567914> (date of access: 26.04.2026).
12. Woszczek H., Wyłomańska A., Chechkin A. Riemann–Liouville fractional Brownian motion with random Hurst exponent. *Chaos: An Interdisciplinary Journal of Nonlinear Science*. 2025. Vol. 35, no. 2. URL: <https://doi.org/10.1063/5.0243975> (date of access: 26.04.2026).
13. Doogue M. P., Polasek T. M. The ABCD of clinical pharmacokinetics. *Therapeutic Advances in Drug Safety*. 2013. Vol. 4, no. 1. P. 5–7. URL: <https://doi.org/10.1177/2042098612469335> (date of access: 26.04.2026).
14. Pereira L. M. Fractal pharmacokinetics. *Computational and Mathematical Methods in Medicine*. 2010. Vol. 11, no. 2. P. 161–184. URL: <https://doi.org/10.1080/17486700903029280> (date of access: 26.04.2026).
15. Dokoumetzidis A., Macheras P. Fractional kinetics in drug absorption and disposition processes. *Journal of Pharmacokinetics and Pharmacodynamics*. 2009. Vol. 36, no. 2. P. 165–178. URL: <https://doi.org/10.1007/s10928-009-9116-x> (date of access: 26.04.2026).
16. Marie N. A pathwise fractional one compartment intra-veinous bolus model. *International Journal of Statistics and Probability*. 2014. Vol. 3, no. 3. URL: <https://doi.org/10.5539/ijsp.v3n3p65> (date of access: 26.04.2026).

17. Fractal characterization of complexity in temporal physiological signals / A. Eke et al. *Physiological Measurement*. 2002. Vol. 23, no. 1. P. R1–R38. URL: <https://doi.org/10.1088/0967-3334/23/1/201> (date of access: 26.04.2026).
18. Esposti F., Ferrario M., Signorini M. G. A blind method for the estimation of the Hurst exponent in time series: Theory and application. *Chaos: An Interdisciplinary Journal of Nonlinear Science*. 2008. Vol. 18, no. 3. P. 033126. URL: <https://doi.org/10.1063/1.2976187> (date of access: 26.04.2026).
19. Lahmiri S. A nonlinear analysis of cardiovascular diseases using multi-scale analysis and generalized hurst exponent. *Healthcare Analytics*. 2023. P. 100142. URL: <https://doi.org/10.1016/j.health.2023.100142> (date of access: 26.04.2026).
20. Typical algorithms for estimating hurst exponent of time sequence: a data analyst's perspective / H.-Y. Zhang et al. *IEEE Access*. 2024. P. 1. URL: <https://doi.org/10.1109/access.2024.3512542> (date of access: 26.04.2026).
21. Mosaic organization of DNA nucleotides / C. K. Peng et al. *Physical Review E*. 1994. Vol. 49, no. 2. P. 1685–1689. URL: <https://doi.org/10.1103/physreve.49.1685> (date of access: 26.04.2026).
22. Matteo T. D., Aste T., Dacorogna M. M. Long-term memories of developed and emerging markets: Using the scaling analysis to characterize their stage of development. *Journal of Banking & Finance*. 2005. Vol. 29, no. 4. P. 827–851. URL: <https://doi.org/10.1016/j.jbankfin.2004.08.004> (date of access: 26.04.2026).
23. Gómez-Águila A., Trinidad-Segovia J. E., Sánchez-Granero M. A. Improvement in Hurst exponent estimation and its application to financial markets. *Financial Innovation*. 2022. Vol. 8, no. 1. URL: <https://doi.org/10.1186/s40854-022-00394-x> (date of access: 26.04.2026).
24. Shimotsu K., Phillips P. C. B. Local Whittle estimation in nonstationary and unit root cases. *The Annals of Statistics*. 2004. Vol. 32, no. 2. P. 656–692.

URL: <https://doi.org/10.1214/009053604000000139> (date of access: 26.04.2026).

25. Exploring stochastic time series structure through wavelet entropy / L. Kirichenko et al. *Lecture Notes on Data Engineering and Communications Technologies*. Cham, 2024. P. 269–283. URL: https://doi.org/10.1007/978-3-031-70959-3_14 (date of access: 26.04.2026).

26. Simonsen I., Hansen A., Nes O. M. Determination of the Hurst exponent by use of wavelet transforms. *Physical Review E*. 1998. Vol. 58, no. 3. P. 2779–2787. URL: <https://doi.org/10.1103/physreve.58.2779> (date of access: 26.04.2026).

27. Trend detection in short time series using discrete wavelet transform / L. Kirichenko et al. *CEUR workshop proceedings*. 2022. Vol. 3347. P. 159–168. URL: https://ceur-ws.org/Vol-3347/Paper_14.pdf (date of access: 28.02.2026).

28. Single-particle diffusion characterization by deep learning / N. Granik et al. *Biophysical Journal*. 2019. Vol. 117, no. 2. P. 185–192. URL: <https://doi.org/10.1016/j.bpj.2019.06.015> (date of access: 26.04.2026).

29. Kowalek P., Loch-Olszewska H., Szwabiński J. Classification of diffusion modes in single-particle tracking data: Feature-based versus deep-learning approach. *Physical Review E*. 2019. Vol. 100, no. 3. URL: <https://doi.org/10.1103/physreve.100.032410> (date of access: 26.04.2026).

30. Single trajectory characterization via machine learning / G. Muñoz-Gil et al. *New Journal of Physics*. 2020. Vol. 22, no. 1. P. 013010. URL: <https://doi.org/10.1088/1367-2630/ab6065> (date of access: 26.04.2026).

31. Argun A., Volpe G., Bo S. Classification, inference and segmentation of anomalous diffusion with recurrent neural networks. *Journal of Physics A: Mathematical and Theoretical*. 2021. Vol. 54, no. 29. P. 294003. URL: <https://doi.org/10.1088/1751-8121/ac070a> (date of access: 26.04.2026).

32. CINNAMON: A hybrid approach to change point detection and parameter estimation in single-particle tracking data / J. Malinowski et al. *Journal*

of *Physics: Photonics*. 2025. URL: <https://doi.org/10.1088/2515-7647/add826> (date of access: 26.04.2026).

33. Differentiable neural network layer for estimating Hurst and anomalous diffusion exponents / R. Lavrynenko et al. *Journal of Physics A: Mathematical and Theoretical*. 2025. URL: <https://doi.org/10.1088/1751-8121/ae17f9> (date of access: 26.04.2026).

34. Likens A. D., Mangalam M., Wong A. Y., Charles A. C., Mills C. Better than DFA? A bayesian method for estimating the hurst exponent in behavioral sciences. *arXiv preprint*. 2023. URL: <https://doi.org/10.48550/arXiv.2301.11262> (дата звернення: 26.04.2026).

35. Whittle P. Estimation and information in stationary time series. *Arkiv för Matematik*. 1953. Vol. 2, no. 5. P. 423–434. URL: <https://doi.org/10.1007/bf02590998> (date of access: 26.04.2026).

36. Ray W. D., Beran J. Statistics for long-memory processes. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*. 1996. Vol. 159, no. 1. P. 180. URL: <https://doi.org/10.2307/2983481> (date of access: 26.04.2026).

37. Fox R., Taqqu M. S. Large-sample properties of parameter estimates for strongly dependent stationary gaussian time series. *The Annals of Statistics*. 1986. Vol. 14, no. 2. P. 517–532. URL: <https://doi.org/10.1214/aos/1176349936> (date of access: 26.04.2026).

38. Robinson P. M. Gaussian Semiparametric estimation of long range dependence. *The Annals of Statistics*. 1995. Vol. 23, no. 5. P. 1630–1661. URL: <https://doi.org/10.1214/aos/1176324317> (date of access: 26.04.2026).

39. Künsch H. Discrimination between monotonic trends and long-range dependence. *Journal of Applied Probability*. 1986. Vol. 23, no. 4. P. 1025–1030. URL: <https://doi.org/10.2307/3214476> (date of access: 26.04.2026).

40. Geweke J., Porter-Hudak S. The estimation and application of long memory time series models. *Journal of Time Series Analysis*. 1983. Vol. 4, no. 4. P. 221–238. URL: <https://doi.org/10.1111/j.1467-9892.1983.tb00371.x> (date of access: 26.04.2026).

41. Hurvich C. M., Deo R., Brodsky J. The mean squared error of Geweke and Porter-Hudak's estimator of the memory parameter of a long-memory time series. *Journal of Time Series Analysis*. 1998. Vol. 19, no. 1. P. 19–46. URL: <https://doi.org/10.1111/1467-9892.00075> (date of access: 26.04.2026).

42. Scaling exponents of time series data: a machine learning approach / S. Raubitzek et al. *Entropy*. 2023. Vol. 25, no. 12. P. 1671. URL: <https://doi.org/10.3390/e25121671> (date of access: 26.04.2026).

43. Andrews D. W. K., Sun Y. Adaptive local polynomial whittle estimation of long-range dependence. *Econometrica*. 2004. Vol. 72, no. 2. P. 569–614. URL: <https://doi.org/10.1111/j.1468-0262.2004.00501.x> (date of access: 26.04.2026).

44. Shimotsu K. Exact Local Whittle estimation of fractional integration with unknown mean and time trend. *Econometric Theory*. 2009. Vol. 26, no. 2. P. 501–540. URL: <https://doi.org/10.1017/s0266466609100075> (date of access: 26.04.2026).

45. Barabási A.-L., Vicsek T. Multifractality of self-affine fractals. *Physical Review A*. 1991. Vol. 44, no. 4. P. 2730–2733. URL: <https://doi.org/10.1103/physreva.44.2730> (date of access: 26.04.2026).

46. Multifractal detrended fluctuation analysis of nonstationary time series / J. W. Kantelhardt et al. *Physica A: Statistical Mechanics and its Applications*. 2002. Vol. 316, no. 1-4. P. 87–114. URL: [https://doi.org/10.1016/s0378-4371\(02\)01383-3](https://doi.org/10.1016/s0378-4371(02)01383-3) (date of access: 26.04.2026).

47. Dynamical generalized Hurst exponent as a tool to monitor unstable periods in financial time series / R. Morales et al. *Physica A: Statistical Mechanics and its Applications*. 2012. Vol. 391, no. 11. P. 3180–3189. URL: <https://doi.org/10.1016/j.physa.2012.01.004> (date of access: 26.04.2026).

48. Barunik J., Kristoufek L. On Hurst exponent estimation under heavy-tailed distributions. *Physica A: Statistical Mechanics and its Applications*. 2010.

Vol. 389, no. 18. P. 3844–3855.

URL: <https://doi.org/10.1016/j.physa.2010.05.025> (date of access: 26.04.2026).

49. Di Matteo T. Multi-scaling in finance. *Quantitative Finance*. 2007. Vol. 7, no. 1. P. 21–36. URL: <https://doi.org/10.1080/14697680600969727> (date of access: 26.04.2026).

50. Taqqu M. S., Teverovsky V., Willinger W. Estimators for long-range dependence: an empirical study. *Fractals*. 1995. Vol. 03, no. 04. P. 785–798. URL: <https://doi.org/10.1142/s0218348x95000692> (date of access: 26.04.2026).

51. Weron R. Estimating long-range dependence: finite sample properties and confidence intervals. *Physica A: Statistical Mechanics and its Applications*. 2002. Vol. 312, no. 1-2. P. 285–299. URL: [https://doi.org/10.1016/s0378-4371\(02\)00961-5](https://doi.org/10.1016/s0378-4371(02)00961-5) (date of access: 26.04.2026).

52. Bryce R. M., Sprague K. B. Revisiting detrended fluctuation analysis. *Scientific Reports*. 2012. Vol. 2, no. 1. URL: <https://doi.org/10.1038/srep00315> (date of access: 26.04.2026).

53. Vestergaard C. L., Blainey P. C., Flyvbjerg H. Optimal estimation of diffusion coefficients from single-particle trajectories. *Physical Review E*. 2014. Vol. 89, no. 2. URL: <https://doi.org/10.1103/physreve.89.022726> (date of access: 26.04.2026).

54. Smirnov N. Table for estimating the goodness of fit of empirical distributions. *The Annals of Mathematical Statistics*. 1948. Vol. 19, no. 2. P. 279–281. URL: <https://doi.org/10.1214/aoms/1177730256> (date of access: 26.04.2026).

55. Massey F. J. The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American Statistical Association*. 1951. Vol. 46, no. 253. P. 68–78. URL: <https://doi.org/10.1080/01621459.1951.10500769> (date of access: 26.04.2026).

56. Stephens M. A. EDF Statistics for goodness of fit and some comparisons. *Journal of the American Statistical Association*. 1974. Vol. 69, no. 347. P. 730–737.

URL: <https://doi.org/10.1080/01621459.1974.10480196> (date of access: 26.04.2026).

57. Burnecki K., Weron A. Fractional Lévy stable motion can model subdiffusive dynamics. *Physical Review E*. 2010. Vol. 82, no. 2. URL: <https://doi.org/10.1103/physreve.82.021130> (date of access: 26.04.2026).

58. Janicki A., Weron A. Stochastic integration. *Simulation and Chaotic Behavior of α -Stable Stochastic Processes*. Boca Raton, 2021. C. 67–109. URL: <https://doi.org/10.1201/9781003208877-4> (дата звернення: 26.04.2026).

59. Sikora G., Burnecki K., Wyłomańska A. Mean-squared-displacement statistical test for fractional Brownian motion. *Physical Review E*. 2017. Vol. 95, no. 3. URL: <https://doi.org/10.1103/physreve.95.032110> (date of access: 26.04.2026).

60. Fractional Brownian motion versus the continuous-time random walk: a simple test for subdiffusive dynamics / M. Magdziarz et al. *Physical Review Letters*. 2009. Vol. 103, no. 18. URL: <https://doi.org/10.1103/physrevlett.103.180602> (date of access: 26.04.2026).

61. Estimating the anomalous diffusion exponent for single particle tracking data with measurement errors - An alternative approach / K. Burnecki et al. *Scientific Reports*. 2015. Vol. 5, no. 1. URL: <https://doi.org/10.1038/srep11306> (date of access: 26.04.2026).

62. Meroz Y., Sokolov I. M. A toolbox for determining subdiffusive mechanisms. *Physics Reports*. 2015. Vol. 573. P. 1–29. URL: <https://doi.org/10.1016/j.physrep.2015.01.002> (date of access: 26.04.2026).

63. Guidelines for the fitting of anomalous diffusion mean square displacement graphs from single particle tracking experiments / E. Kepten et al. *PLOS ONE*. 2015. Vol. 10, no. 2. P. e0117722. URL: <https://doi.org/10.1371/journal.pone.0117722> (date of access: 26.04.2026).

64. Ensemble-based correction for anomalous diffusion exponent estimation in single-particle tracking / R. Lavrynenko et al. *Applied sciences*. 2025. Vol. 15, no. 14. URL: <https://doi.org/10.3390/app15148000> (date of access: 28.02.2026).

65. Shergin V., Chala L., Udovenko S. Statistical properties of the Hurst exponent estimates for fractional Levy motion. *2020 IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET)*, Lviv-Slavske, Ukraine, 25–29 February 2020. 2020. URL: <https://doi.org/10.1109/tcset49122.2020.235421> (date of access: 26.04.2026).

66. Shergin V. Estimating the Hurst exponent of fractional Levy motion by the fractional moments method. *2019 IEEE International Scientific-Practical Conference Problems of Infocommunications, Science and Technology (PIC S&T)*, Kyiv, Ukraine, 8–11 October 2019. 2019. URL: <https://doi.org/10.1109/picst47496.2019.9061298> (date of access: 26.04.2026).

67. Seckler H., Metzler R. Bayesian deep learning for error estimation in the analysis of anomalous diffusion. *Nature Communications*. 2022. Vol. 13, no. 1. URL: <https://doi.org/10.1038/s41467-022-34305-6> (date of access: 26.04.2026).

68. Terziyan V., Vitko O., Terziyan O. Semantic AI for future industries: bridging explainability and integration in black box models. *Procedia Computer Science*. 2026. Vol. 277. P. 2235–2246. URL: <https://doi.org/10.1016/j.procs.2026.02.261> (date of access: 26.04.2026).

69. Non-Gaussian, non-ergodic, and non-Fickian diffusion of tracers in mucin hydrogels / A. G. Cherstvy et al. *Soft Matter*. 2019. Vol. 15, no. 12. P. 2526–2551. URL: <https://doi.org/10.1039/c8sm02096e> (date of access: 26.04.2026).

70. Parameter estimation of long memory stochastic processes with deep neural networks / B. Csanády et al. *Frontiers in Artificial Intelligence and*

Applications. 2024. URL: <https://doi.org/10.3233/faia240784> (date of access: 25.04.2026).

71. Wang Z., Yan W., Oates T. Time series classification from scratch with deep neural networks: A strong baseline. *2017 International Joint Conference on Neural Networks (IJCNN)*, Anchorage, AK, USA, 14–19 May 2017. 2017. URL: <https://doi.org/10.1109/ijcnn.2017.7966039> (date of access: 26.04.2026).

72. Convolutional hyper basis function neural network and its online learning in the image recognition task / Y. Bodyanskiy et al. *CEUR workshop proceedings*. 2021. Vol. 2853. P. 108–117. URL: <https://ceur-ws.org/Vol-2853/paper8.pdf> (date of access: 28.02.2026).

73. Towards length-versatile and noise-robust radio frequency fingerprint identification / G. Shen et al. *IEEE Transactions on Information Forensics and Security*. 2023. P. 1. URL: <https://doi.org/10.1109/tifs.2023.3266626> (date of access: 25.04.2026).

74. Di Persio L., Dufera T. T. Deep neural network model for Hurst exponent: learning from R/S analysis. *Mathematics*. 2024. Vol. 12, no. 22. P. 3483. URL: <https://doi.org/10.3390/math12223483> (date of access: 25.04.2026).

75. Lin M., Chen Q., Yan S. Network In Network. *arXiv*. 2013. URL: <https://doi.org/10.48550/arXiv.1312.4400> (date of access: 26.04.2026).

76. Dempster A., Schmidt D. F., Webb G. I. Hydra: competing convolutional kernels for fast and accurate time series classification. *Data Mining and Knowledge Discovery*. 2023. URL: <https://doi.org/10.1007/s10618-023-00939-3> (date of access: 25.04.2026).

77. Online feature vector restoration in data stream mining tasks / V. Kartashov et al. *CEUR workshop proceedings*. 2025. Vol. 4164. P. 227–239. URL: <https://ceur-ws.org/Vol-4164/paper15.pdf> (date of access: 28.02.2026).

78. A fault diagnosis scheme using Hurst exponent for metal particle faults in GIL/GIS / D. Duan et al. *Sensors*. 2022. Vol. 22, no. 3. P. 862. URL: <https://doi.org/10.3390/s22030862> (date of access: 25.04.2026).

79. Impact of outliers on median-based and mean-based methods for anomalous diffusion exponent estimation / R. Lavrynenko et al. *Journal of Physics: Photonics*. 2025. URL: <https://doi.org/10.1088/2515-7647/adc04e> (date of access: 26.04.2026).

80. Short-duration monofractal signals for heart failure characterization using CNN-ELM models / J. L. López et al. *Applied Sciences*. 2025. Vol. 15, no. 21. P. 11453. URL: <https://doi.org/10.3390/app152111453> (date of access: 25.04.2026).

81. Rapid detection of fetal compromise using input length invariant deep learning on fetal heart rate signals / L. Mendis et al. *Scientific Reports*. 2024. Vol. 14, no. 1. URL: <https://doi.org/10.1038/s41598-024-63108-6> (date of access: 25.04.2026).

82. Deep learning the Hurst parameter of linear fractional processes and assessing its reliability / D. Boros et al. *Quality and Reliability Engineering International*. 2024. URL: <https://doi.org/10.1002/qre.3641> (date of access: 25.04.2026).

83. GitHub - aielte-research/LMSParEst: Parameter estimation of long memory stochastic processes with deep neural networks / B. Csanády et al. GitHub. URL: <https://github.com/aielte-research/LMSParEst> (date of access: 26.06.2026).

84. AnDi 2 Benchmark dataset / M.-G. Gorka et al. Zenodo. URL: <https://zenodo.org/doi/10.5281/zenodo.14281479> (date of access: 26.06.2026).

85. Рябова Н., Лавриненко С., Лавриненко Р. Дослідження впливу наявності меж при симуляції фрактального броунівського руху на точність оцінки показника Херста. *Радіoeлектроніка та молодь у XXI столітті. Т. 6 : Конференція "Інформаційні інтелектуальні системи"*. Харків, Україна, 2024. URL: <https://doi.org/10.30837/iyf.iis.2024.091> (дата звернення: 15.05.2026).

86. Statistical methods for investigation of heterogeneous diffusion in single-particle tracking / L. Kirichenko et al. *2024 14th International Conference*

on Dependable Systems, Services and Technologies (DESSERT), Athens, Greece, 11–13 October 2024. 2024. P. 1–6. URL: <https://doi.org/10.1109/dessert65323.2024.11122178> (date of access: 15.05.2026).

87. Lavrynenko S. GitHub - SophiaLavr/conv1d-hurst-estimation-fbm: Benchmark of classical and Conv1D-based Hurst exponent estimators for fractional Brownian motion and fractional Gaussian noise. GitHub. URL: <https://github.com/SophiaLavr/conv1d-hurst-estimation-fbm> (date of access: 26.06.2026)..

88. Natu M., Bachute M., Kotecha K. HCLA_CBiGRU: Hybrid convolutional bidirectional GRU based model for epileptic seizure detection. *Neuroscience Informatics*. 2023. P. 100135. URL: <https://doi.org/10.1016/j.neuri.2023.100135> (date of access: 15.05.2026).

89. A channel selection MLP-Mixer network for EEG-based motor imagery BCI / H. Wang et al. *Biomedical Signal Processing and Control*. 2026. Vol. 112. P. 108715. URL: <https://doi.org/10.1016/j.bspc.2025.108715> (date of access: 15.05.2026).