МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

ЛЬВІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ІВАНА ФРАНКА

Факультет прикладної математики та інформатики
(повне найменування назва факультету)

Кафедра інформаційних систем
(повна назва кафедри)

Магістерська робота

Інформаційна технологія прогнозування поширення COVID-19 з використанням мультифрактального аналізу

Виконав: студент групи ПМІМ-22с
спеціальності
122 - комп'ютерні науки
(шифр і назва спеціальності)

(підпис) Кухар О.Є.
(прізвище та ініціали)

Керівник (підпис) Соколовський. Я.І
(прізвище та ініціали)

Рецензент (підпис) Кособуцький П.С.
(прізвище та ініціали)

Львів – 2018

ЛЬВІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ІВАНА ФРАНКА

Факультет *прикладної математики та інформатики*
Кафедра *інформаційних систем*
Спеціальність 122 – комп'ютерні науки
(шифр і назва)

«ЗАТВЕРДЖУЮ»
Завідувач кафедри *проф. Шинкаренко Г.А.*
" *5* " *вересня* **20** *22* **року**

З А В Д А Н Н Я

НА МАГІСТЕРСЬКУ РОБОТУ СТУДЕНТУ

Кухару Олександру Євгеновичу
(прізвище, ім'я, по батькові)

1 .Тема роботи *Інформаційна технологія прогнозування поширення COVID-19 з використанням мультифрактального аналізу*

керівник роботи *Соколовський Ярослав Іванович, доктор технічних наук, професор,*
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)
затверджені Вченою радою факультету від " 13 " *вересня* 20 *22* року № 15

2. Строк подання студентом роботи *12.12.2022 р.*

3. Вихідні дані до роботи *Література та інтернет-ресурси за тематикою роботи, датасет захворюваності на COVID-19*

4. Зміст магістерської роботи (перелік питань, які потрібно розробити)

1. Аналіз існуючого стану
2. Інформаційне забезпечення
3. Математичне та алгоритмічне забезпечення
4. Програмне забезпечення

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень)
Графіки, презентація дипломної роботи

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв

7. Дата видачі завдання *05.09.2022 р.*

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів магістерської роботи	Строк виконання етапів роботи	Примітка
1.	*Аналіз існуючого стану*	*Вересень*	*Виконано*
2.	*Пошук датасету*	*Жовтень*	*Виконано*
3.	*Розробка алгоритмів*	*Листопад*	*Виконано*
4.	*Програмна реалізація*	*Листопад*	*Виконано*
5.	*Оформлення роботи*	*Грудень*	*Виконано*

Студент _____ *Кухар О. Є.*
(підпис) (прізвище та ініціали)

Керівник роботи _____ *проф. Соколовський Я. І*
(підпис) (прізвище та ініціали)

ЗМІСТ

ЗМІСТ 4
ТЕХНІЧНЕ ЗАВДАННЯ 5
АНОТАЦІЯ 6
ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ 7
ВСТУП 8
РОЗДІЛ 1. АНАЛІЗ ІСНУЮЧОГО СТАНУ 10
1.1. Аналіз предметної області. 10
1.2. Аналіз моделей прогнозування. 10
1.3. Аналіз програмних засобів. 11
1.4. Висновки. 11
РОЗДІЛ 2. ІНФОРМАЦІЙНЕ ЗАБЕЗПЕЧЕННЯ 12
2.1. Датасет для прогнозування поширення Covid-19. 12
2.2. Дані по захворюваності, нових випадках і смертях у Франції. 13
2.3. Висновки. 13
РОЗДІЛ 3. МАТЕМАТИЧНЕ ТА АЛГОРИТМІЧНЕ ЗАБЕЗПЕЧЕННЯ. 14
3.1. Характеристики часових рядів. 14
3.2. Мультифрактальний аналіз. 16
3.3. Модель ARFIMA. 18
3.4. Висновки. 19
РОЗДІЛ 4. ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ 20
4.1. Програмна реалізація мультифрактального аналізу. 20
4.2. Програмна реалізація моделі ARFIMA. 29
4.3. Результати досліджень. 29
4.4. Аналіз існуючих результатів. 31
4.5. Висновки. 38
ВИСНОВКИ 39
ДОДАТКИ 40
Додаток А. Реалізація знаходження локального RMS 40
Додаток Б. Реалізація алгоритму DFA. 42
Додаток В. Реалізація алгоритму MFDFA. 43
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ 45

ТЕХНІЧНЕ ЗАВДАННЯ

1. Підготувати набір даних по захворюваності від COVID-19. Перевірити набір даних на актуальність та наявність аномалій, при потребі провести чистку даних для подальшого використання.
2. Проаналізувати та дослідити алгоритми мультифрактального аналізу. Вибрати алгоритм мультифрактального аналіз набору даних.
3. Обґрунтувати вибір моделі прогнозування та параметрів.
4. Розробити програмне забезпечення для реалізації алгоритмів мультифрактального аналізу та прогнозування. Провести мультифрактальний аналіз датасету.
5. Розробити модель прогнозування. Провести порівняння результатів прогнозування з іншими наявними моделями.

АНОТАЦІЯ

Магістерська робота студента Кухара Олександра Євгеновича на тему: “Інформаційна технологія прогнозування поширення COVID-19 з використанням мультифрактального аналізу” містить 37 сторінок, 10 джерел за списком використаної літератури та розділ додатків.

У магістерській роботі було створено систему на мові програмування R для моделювання та прогнозування поширення COVID-19 у Франції з використанням фрактальної моделі ARFIMA. Створено систему для проведення мультифрактального аналізу за допомогою алгоритму MFDFA.

Ключові слова: часовий ряд, MFDFA, ARFIMA, мультифрактальний аналіз, довгострокова залежність, DFA, R language, Python.

The master thesis of student Oleksandr Kukhar on the topic: “Information system for prediction of COVID-19 spread using multifractal analysis” contains 37 pages, 10 sources from the list of references and a section of appendies.

In the master thesis there was created a system in the R programming language for modeling and forecasting the spread of COVID-19 using the fractal model ARFIMA. Created a system for multifractal analysis with MFDFA algorithm.

Keywords: time series, MFDFA, ARFIMA, multifractal analysis, long-range dependence, DFA, R language, Python.

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

AR	Autoregressive model (авторегресійна модель).
MA	Moving average model (модель з ковзним середнім).
ARFIMA	Autoregressive fractionally integrated moving average (авторегресійна дробово-інтегрована модель з ковзним середнім).
ACF	Autocorrelation function (Автокореляційна функція).
PACF	Partial autocorrelation function (Функція часткової автореляції).
RMS	Кореневе середнє квадратичне .
DFA	Метод флуктуаційного аналізу.
MFDFA	Метод мультифрактального флуктуаційного аналізу.
RMSE	Кореневе середньоквадратичне відхилення.
MAPE	Середня абсолютна відсоткова помилка.
Датасет	Набір даних.

ВСТУП

Актуальність завдання. Перші випадки невідомої хвороби появились у грудні 2019 року у Китаї. Через місяць 7 січня 2020 року цей вірус ідентифікували як COVID-19. Були запровадженні чисельні обмеження для громадян Китаю з метою зупинити поширення вірусу, проте незважаючи на всі зусилля вірус поширився у всі частини світу. З часом пандемія почала втрачати свій фокус уваги, а з початком війни в Україні практично зникла з інформаційного поля. Хоч і інформаційний простір на даний момент заповнений новинами про війну в Україні і її наслідками для всього світу, вірус не припинив своє існування і його нові спалахи починають виникати знову.

Мета роботи. Під час поширення будь-якого вірусу гостро виникає питання подолання його наслідків. Проблема полягає у тому, що лікарні містять обмежені ресурси лікування хворих і вони зазвичай не розраховані на кількості які потрібно лікувати під час таких пандемій. У такому випадку потрібно вживати заходи для сповільнення поширення вірусу, щоб не допустити ситуацій, коли кількість хворих людей перевищує максимальну кількість людей, яких можуть прийняти лікарні. Тому метою цієї роботи є розробка інформаційної системи, яка би могла передбачувати подальше поширення вірусу по світу або певній країні. Для досягнення мети були виконані завдання, зазначені у розділі «Технічне завдання».

Об’єкт дослідження. Динаміка поширення вірусу, яку необхідно перетворити у часовий ряд, використавши попередньо знайдені дані. Дані представлені у вигляді таблиць, де щоденно описана кількість нових випадків захворювання у певній країні.

Предмет дослідження. Інформаційна технологія яка буде проводити аналіз поширення вірусу та складати прогнози.

Практичне значення роботи. Систему можна використовувати для визначення поточного стану ситуації, визначення кількості нових випадків захворювань якщо її не змінювати і відповідно чи необхідно приймати певні дії для зміни цієї ситуації.

Наукова новизна. Використання алгоритму мультифрактального аналізу MFDFA на даних захворюваності COVID-19, а також використання фрактальної моделі ARFIMA для складання прогнозів захворюваності.

РОЗДІЛ 1. АНАЛІЗ ІСНУЮЧОГО СТАНУ

1.1. Аналіз предметної області.

Поточна пандемія COVID-19 була спричинена коронавірусом типу SARS CoV-2. Коронавіруси (CoVs) – це велика група вірусів, багато з яких спричиняють респіраторні захворювання починаючи з простої застуди та закінчуючи *важким гострим респіраторним синдромом* (SARS) і *близькосхідним коронавірусним респіраторним синдромом* (MERS), які мають високі показники смертності та були вперше виявлені у 2003 і 2012 роках відповідно.

Станом на сьогодні, зоонозне джерело SARS-CoV-2 не визначене. Перші випадки появи вірусу і людей були задокументовані у місті Ухань, Китай у грудні 2019 року. Однією з теорій виступає те, що вірус міг поширитись до людини від кажанів через іншого хоста, так як зазвичай люди дуже рідко контактують з кажанами. Таким проміжним хостом може бути як дика так і свійська тварина. У будь-якому випадку, допоки не буде знайдено і усунено джерело вірусу, існує ризик подальших спалахів коронавірусу у людей, таких які ми зараз переживаємо [1].

1.2. Аналіз моделей прогнозування.

Моделі прогнозування працюють з часовими рядами. Часовими рядами називають впорядковану послідовність точок даних. При їх аналізі визначаються певні закономірності, на основі яких будуються майбутні прогнози. Існують різні моделі прогнозування, які призначені для часових рядів з різними закономірностями. Серед них основними є:

- AR – авторегресійні моделі.
- I – інтегровані моделі.
- MA – моделі з ковзним середнім.

Ці три моделі також можуть бути скомбіновані між собою. Комбінаціями цих моделей є ARMA (авторегресійні моделі з ковзним середнім) і ARIMA (авторегресійні інтегровані моделі з ковзним середнім). Продовженням цих моделей є ARFIMA (авторегресійні дробово-інтегровані моделі з ковзним середнім). Також існують моделі які працюють з багатовимірними часовими рядами, тобто рядами в яких є кілька змінних. Такі моделі містять приставку V (vector) у назві. Наприклад: VAR або VARMA.

Враховуючи особливості часового ряду, які будуть описані пізніше, найкращою моделлю для його прогнозування буде ARFIMA.

1.3. Аналіз програмних засобів.

Для виконання досліджень було використано мову програмування R версії 4.2.0. Ця мова спеціалізується на обробці, аналізі і зображені статистичних даних. Також було використано бібліотеки arfima версії 1.8.0 для створення однойменної моделі та MFDFA версії 1.1 для аналізу датасету. Для проведення попередньої обробки набору даних та зображення проміжних результатів роботи алгоритмів використовувалась мова програмування Python версії 3.6.5, з використанням бібліотек matplotlib, statistics, та numpy.

1.4. Висновки.

Коронавіруси є великою групою вірусів, які можуть спричинити респіраторні захворювання різної важкості. Поточна пандемія була спричинена через коронавірус типу SARS CoV-2, джерело якого на даний момент не визначене. Через це він може спалахувати знову і знову. Для його моделювання і прогнозування, серед наявних моделей було вибрано фрактальну модель ARFIMA. Програмна реалізація відбувалась на мові програмування R, з використанням бібліотек arfima та MFDFA.

РОЗДІЛ 2. ІНФОРМАЦІЙНЕ ЗАБЕЗПЕЧЕННЯ

2.1. Датасет для прогнозування поширення Covid-19.

Датасет був взятий з ресурсу European Centre for Disease Prevention and Control [8]. У ньому містяться дані про кількість випадків захворювання, кількість смертей та кількість вилікуваних людей у Європі. Ці дані розбиті по країнам.

dateRep	day	month	year	cases	deaths	countriesAndTerritories	geoId	countryterritoryCode	popData2020	continentExp
23.10.2022	23	10	2022	3557	0	Austria	AT	AUT	8901064	Europe
22.10.2022	22	10	2022	5494	4	Austria	AT	AUT	8901064	Europe
21.10.2022	21	10	2022	7776	4	Austria	AT	AUT	8901064	Europe
20.10.2022	20	10	2022	8221	6	Austria	AT	AUT	8901064	Europe
19.10.2022	19	10	2022	10007	8	Austria	AT	AUT	8901064	Europe
18.10.2022	18	10	2022	13204	7	Austria	AT	AUT	8901064	Europe
17.10.2022	17	10	2022	9964	8	Austria	AT	AUT	8901064	Europe
16.10.2022	16	10	2022	6606	12	Austria	AT	AUT	8901064	Europe
15.10.2022	15	10	2022	8818	6	Austria	AT	AUT	8901064	Europe
14.10.2022	14	10	2022	11751	10	Austria	AT	AUT	8901064	Europe
13.10.2022	13	10	2022	13068	14	Austria	AT	AUT	8901064	Europe
12.10.2022	12	10	2022	14305	13	Austria	AT	AUT	8901064	Europe
11.10.2022	11	10	2022	18498	11	Austria	AT	AUT	8901064	Europe
10.10.2022	10	10	2022	13369	10	Austria	AT	AUT	8901064	Europe
09.10.2022	9	10	2022	8689	10	Austria	AT	AUT	8901064	Europe
08.10.2022	8	10	2022	11230	12	Austria	AT	AUT	8901064	Europe
07.10.2022	7	10	2022	14391	18	Austria	AT	AUT	8901064	Europe
06.10.2022	6	10	2022	15560	10	Austria	AT	AUT	8901064	Europe
05.10.2022	5	10	2022	15291	13	Austria	AT	AUT	8901064	Europe
04.10.2022	4	10	2022	18531	16	Austria	AT	AUT	8901064	Europe
03.10.2022	3	10	2022	14200	16	Austria	AT	AUT	8901064	Europe
02.10.2022	2	10	2022	8525	4	Austria	AT	AUT	8901064	Europe
01.10.2022	1	10	2022	10344	4	Austria	AT	AUT	8901064	Europe
30.09.2022	30	9	2022	12571	9	Austria	AT	AUT	8901064	Europe
29.09.2022	29	9	2022	13061	5	Austria	AT	AUT	8901064	Europe
28.09.2022	28	9	2022	12747	6	Austria	AT	AUT	8901064	Europe
27.09.2022	27	9	2022	15316	5	Austria	AT	AUT	8901064	Europe
26.09.2022	26	9	2022	10927	9	Austria	AT	AUT	8901064	Europe
25.09.2022	25	9	2022	6088	6	Austria	AT	AUT	8901064	Europe
24.09.2022	24	9	2022	6931	3	Austria	AT	AUT	8901064	Europe
23.09.2022	23	9	2022	8564	4	Austria	AT	AUT	8901064	Europe
22.09.2022	22	9	2022	7940	3	Austria	AT	AUT	8901064	Europe

Рис. 2.1. Приклад даних з датасету з даними про захворювання.

Перші записи починаються з 26 січня 2020 року. Всього у датасеті налічується 1000 записів по більшості країн. Для своїх досліджень я вирішив взяти тільки дані про випадки нових захворювань і смертей у Франції.

2.2. Дані по захворюваності, нових випадках і смертях у Франції.

Після вибору країни датасет потрібно привести до вигляду часового ряду. При детальному розгляді графіків на рис. 2.2. можна побачити, що вірус поширювався нерівномірно. Спочатку захворюваність різко піднімається, а потім починає затримуватись. Також на графіку з новими випадками COVID-19 видно дрібні флуктуації. При детальному перегляді було обчислено, що ці флуктуації трапляються один раз у 6-7 днів. З цього можна зробити висновок, що вони можуть бути пов’язані з малою кількістю тестувань у період з суботи по неділю.

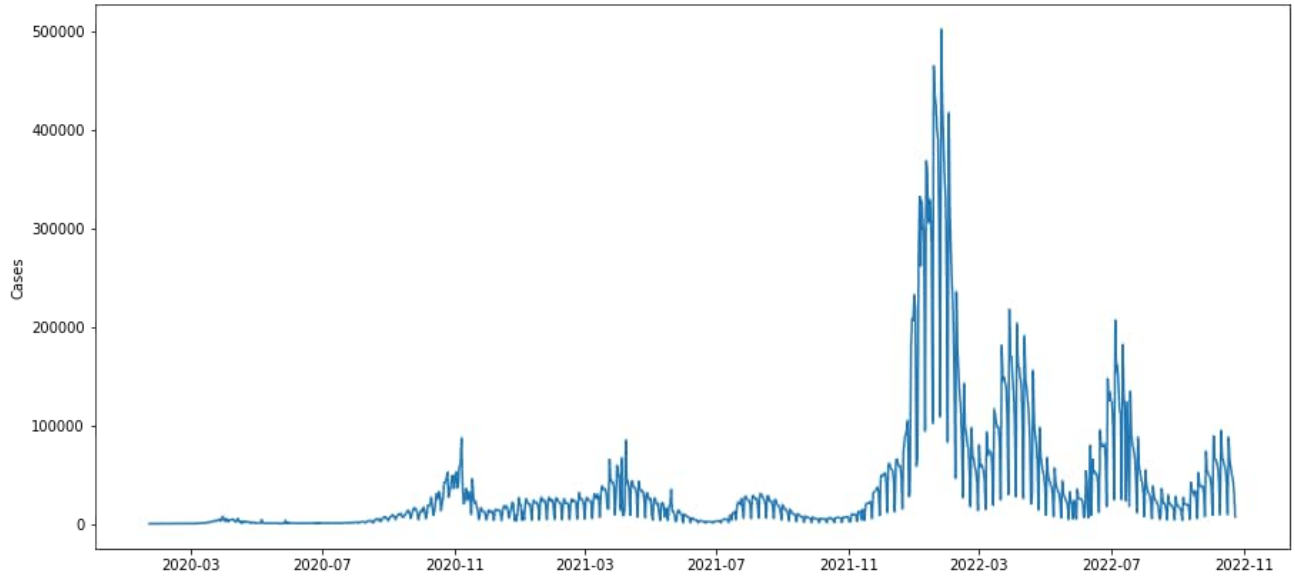

Рис. 2.2. Графік кількості нових випадків поширення COVID-19 у Франції.

2.3. Висновки.

Для проведення досліджень було обрано набір даних, який містив 1000 записів, починаючи з кінця січня 2020 року і закінчуючи жовтнем 2022 року. Серед доступних країн було вибрано Францію. Після перетворення набору даних до виду часового ряду та детального огляду було виявлено певні періодичності та флуктуації.

РОЗДІЛ 3. МАТЕМАТИЧНЕ ТА АЛГОРИТМІЧНЕ ЗАБЕЗПЕЧЕННЯ.

3.1. Характеристики часових рядів.

При аналізі часових рядів традиційно припускають, що зв’язок між значеннями швидко зникає зі збільшенням часової відстані між ними. Довгострокова залежність (Long-range dependence), або довга пам’ять – це властивість, яка може виникати при аналізі певних часових рядів. Довгострокову залежність вперше описав британський гідролог Гарольд Едвін Херст у своїй роботі при дослідженні гідрологічних даних. Після цього у всіх інших статистиках почали використовувати довгострокову пам’ять, яку характеризують критерієм Херста, що вказує на сильний зв’язок між значеннями часового ряду, які знаходяться на різних часових відстанях. [3]

Для створення моделі та прогнозування часових рядів потрібно спочатку визначити певні характеристики, а саме наявність тренду, сезонності та довгої пам’яті. З графіку на рис. 2.2. ми могли побачити просідання кількості нових захворювань кожні 6-7 днів. Це явище називається сезонністю. Також ми могли побачити, що вірус поширювався так званими “хвилями”. Ці хвилі також відносяться до сезонних явищ. Загальне збільшення кількості нових випадків захворювань є трендом.

Автокореляційна функція (Autocorrelation function ACF) – це один з корисних методів для знаходження тренду і періодичності даних, позволяючи проводити візуальні спостереження. Також за допомогою автокореляційної функції можна визначити властивості довгої пам’яті.

На рис. 3.1. показані графіки автокореляції та часткової автокореляції.

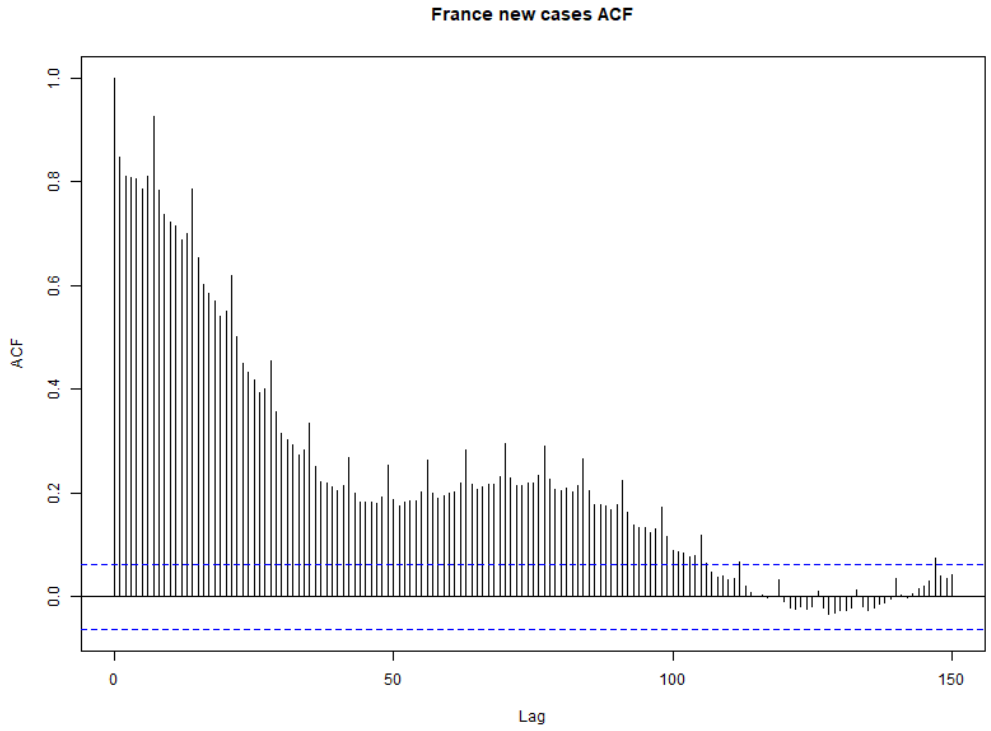

Рис. 3.2. Графік автокореляції для нових випадків Covid-19 у Франції.

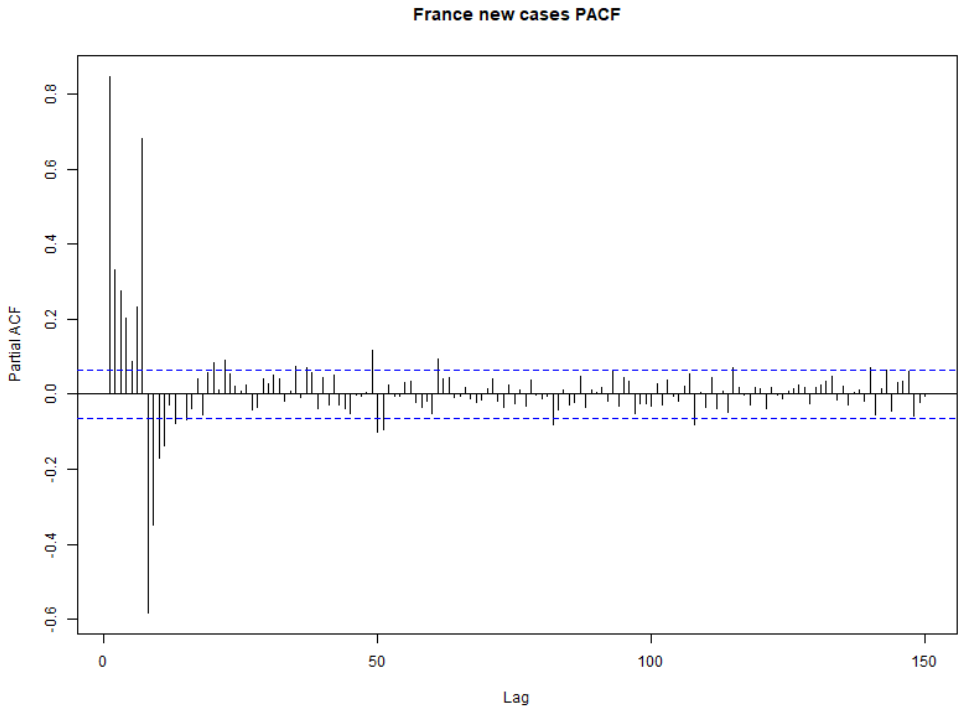

Рис. 3.3. Графік часткової автокореляції для нових випадків Covid-19 у Франції.

Перший графік показує наскільки статистично залежним є значення часового ряду від його попередніх значень. Другий показує те саме, але з врахуванням впливу значень між ними. У нашому випадку потрібно 120 часових затримок, щоб знизити ефект залежності до статистично незначущого рівня. А це вказує на наявність довгої пам’яті у нашій системі.

3.2. Мультифрактальний аналіз.

Теорія фракталів і мультифракталів широко використовується для опису властивостей самоподібності і складного скейлингу, яке можна побачити у різних фізичних ситуаціях. До фракталів відносять геометричні об’єкти які мають сильно зрізану форму і містять певну повторюваність у широкому діапазоні масштабів. Повторюваність може бути повною (в такому випадку ми маємо справу з регулярними фракталами), або може існувати певний елемент випадковості (такі фрактали називають випадковими). Структура випадкових фракталів на малих масштабах не є ідентичною структурі всього об’єкту, але їх статистичні характеристики співпадають і властивості самоподібності зберігаються після усереднення по статистично незалежним реалізаціям об’єкту.

Для кількісного опису фракталу достатньо одної величини – розмірності Хаусфорда або показника скейлингу, який описує збереження геометрії або статистичних характеристик при зміні масштабу. Але існує багато явищ, які потребують використання поняття фракталу на складні структури з більш ніж одним показником скейлингу. Такі структури часто характеризуються цілим спектром показників і розмірність Хаусфорда є лише одним з них. Складні фрактали або мультифрактали найчастіше зустрічаються в природі, так як прості самоподібні об’єкти зазвичай є ідеалізацією певних явищ. Мультифрактальний підхід означає, що об’єкт який вивчають можна розділити на частини, кожна з яких має власні властивості самоподібності. Даний підхід успішно використовувався у різних областях: при вивчені агрегаційних

властивостей клітинних елементів крові у біології, для опису руйнування матеріалів в теорії розвинутої гідродинамічної турбулентності і т.д. Різні об’єкти можна віднести до спеціального класу мультифракталів і доволі складно знайти область науки де не зустрічались би представники цього класу.

В рамках мультифрактального підходу розглядаються так звані часткові суми (або узагальнені статистичні суми)

$$z_q(\varepsilon) = \sum_{i=}^{N(\varepsilon)} \mu_i^q(\varepsilon), \quad (1)$$

де $N(\varepsilon)$ – число елементів покриття ε розміром досліджуваної множини A, μ_i – міра і-го елемента елементу покриття, $q \in R$. Залежність функцій Z від ε, як правило містить степеневий характер типу

$$Z(q,\varepsilon) \sim \varepsilon^{(q-1)Dq}, \quad (2)$$

де величини D_q є *узагальненими фрактальними розмірностями.* Коефіцієнт q – 1 включений в показник степеню для автоматичного виконання рівняння Z (1, г) = 1, яке означає нормування міри μ. Зазвичай виконується дане позначення:

$$\tau(q) = qD_q - 1, \quad (3)$$

Величини $\tau(q)$ називають скейлінговими експонентами або масовим показником.

Для простих фракталів D_q = const. В загальному випадку мультифрактальних об’єктів значення D_q монотонно зменшуються зі зростанням q. Це зменшення можна розглядати в якості діагностичного критерію, який підтверджує, що аналізований об’єкт є неоднорідним фракталом. Внаслідок чого, залежність $\tau(q)$ буде прямою лінією для однорідних фракталів і нелінійною функцією для неодонорідних.

В рамках цього методу функції $f(a)$ знаходять за допомогою перетворення Лежандра

$$\begin{cases} \alpha = \frac{d\tau}{dq} \\ f(\alpha) = q\alpha - \tau(q) \end{cases}, \quad (4)$$

Таким чином, по відомому спектру скейлінгових експонент $\tau(q)$ можна одразу визначити шукану функцію $f(\alpha)$ [2].

3.3. Модель ARFIMA.

У зв'язку зі збільшенням попиту на моделювання та аналізу довгострокової залежності і самоподібності у часових рядах, таких як фінансові дані, дані про трафік мереж зв'язку і підводного шуму, методи обробки сигналу дробового порядку набули великої популярності.

У порівнянні з попередніми моделями, модель ARFIMA показує кращу підгонку і результат, коли працює з даними які мають властивості довгої пам'яті.

Модель ARFIMA є комбінацією моделей AR і MA з деякими доповненнями де:

Autoregressive model (AR):

$$X_t = c + \sum_{i=1}^{p} \varphi_i X_{t-i} + \varepsilon_t, \quad (5)$$

Де $\varphi_1, \dots, \varphi_p$ – авторегресивні параметри, c – стала і випадкова величина ε_t – білий шум. Деякі параметри мають обмеження, щоб модель залишалась стаціонарною. У статистиці і обробці сигналів авторегресивна модель використовується для репрезентації типу випадкового процесу, наприклад опису певних процесів які змінюються з часом у природі, економіці і т.д.

Moving average model (MA):

$$X_t = \mu + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t, \quad (6)$$

Де $\theta_1, \dots, \theta_q$ – параметри ковзного середнього моделі, μ – очікування від X_t(Часто використовують значення 0) і $\varepsilon_t, \varepsilon_{t-1}, \dots$ - білий шум. Ковзне середнє згладжує часовий ряд, що може привести до циклічних графіків, або таких які

показують загальну тенденцію навіть коли початкові дані є незалежними випадковими величинами зі сталим середнім. Ця особливість зменшує користь даної моделі.

Описані вище AR і MA моделі можуть бути узагальнені до виду:

Autoregressive integrated moving average (ARIMA):

$$(1 - \sum_{i=1}^{p} \varphi_i B^i)(1 - \mathrm{B})^d (X_t - \mu) = (1 + \sum_{i=1}^{q} \theta_i B^i)\varepsilon_t, \quad (7)$$

$(1 - \mathrm{B})^d$ називають оператором диференціювання ∇^d. ARMA і ARIMA моделі можуть фіксувати тільки властивості короткострокової залежності, так як для параметру d використовуються тільки цілі числа. Тому для того, щоб фіксувати властивості довгої пам'яті у фрактальних системах, було представлено модель ARFIMA(p,d,q). Оператор диференціювання можна представити для кожного дійсного d за допомогою Гамма функції [3]:

$$(1 - B)^d = \sum_{k=0}^{\infty} \binom{d}{k} (-B)^k = \sum_{k=0}^{\infty} \frac{\Gamma(d+1)}{\Gamma(k+1)\Gamma(d+1-k)} (-B)^k, \quad (8)$$

$$\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt, \quad (9)$$

3.4. Висновки.

Однією з основних характеристик часових рядів є наявність довгострокової залежності. Вона може вимірюватись критерієм Херста. Також використовуються графіки функцій автокореляції та часткової автокореляцій. Під час їх аналізу було визначено, що у даному часовому ряді потрібно 120 часових затримок для того, щоб звести ефект залежності до статистично незначущого рівня. Але для опису певних часових рядів такого аналізу може бути недостатньо. Для цього використовують методи мультифрактального аналізу, які можуть розділити часовий ряд на частини, кожна з яких має власні властивості самоподібності. Для складання прогнозів використовуються різні моделі. Фрактальна модель ARFIMA, яка є комбінацією моделей AR і MA, показує кращі результати прогнозування коли працює з даними, які містять довгу пам'ять.

РОЗДІЛ 4. ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ

4.1. Програмна реалізація мультифрактального аналізу.

Для проведення мультифрактального аналізу було використано бібліотеку MFDFA, яка містить такі методи, як MultiFractal Detrended Fluctuation Analysis (MFDFA), MultiFractal Detrended Cross-Correlation Analysis (MFXDFA) I Multiscale Multifractal Analysis (MMA).

Функція приймає rsx – вектор елементів часового ряду, scale – вектор шкал, m – число поліноміального порядку і q – порядок моменту. Після проведення обчислень функція поверне Hq – критерій Херста, tau_q – критерій Масса, spec – мультифрактальний спектр та Fq – функція флуктуації.

Спочатку алгоритм проводить певні фільтрації для того, щоб привести часовий ряд з типу випадкового шуму до часового ряду виду випадкового блукання. Для цього спочатку від кожного значення часового ряду віднімається середнє значення, а потім обчислюється кумулятивна сума для кожного елементу.

$$Y(i) = \sum_{n=0}^{i} x(i) - x, \tag{10}$$

У результаті часовий ряд набуває такого вигляду:

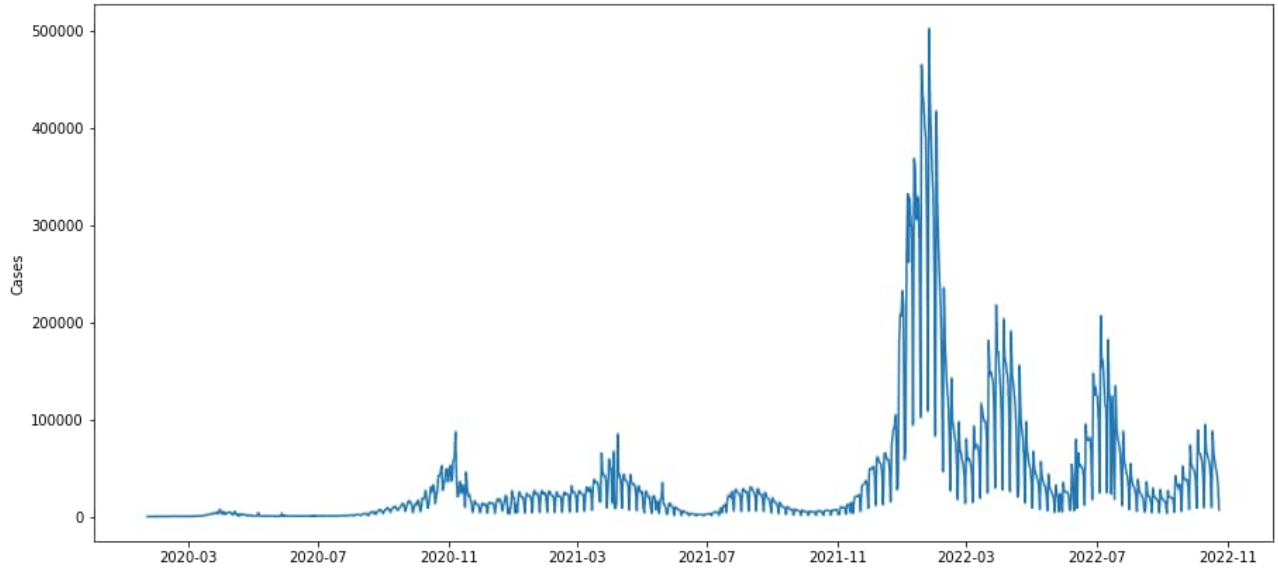

Рис. 4.1. Часовий ряд до перетворень.

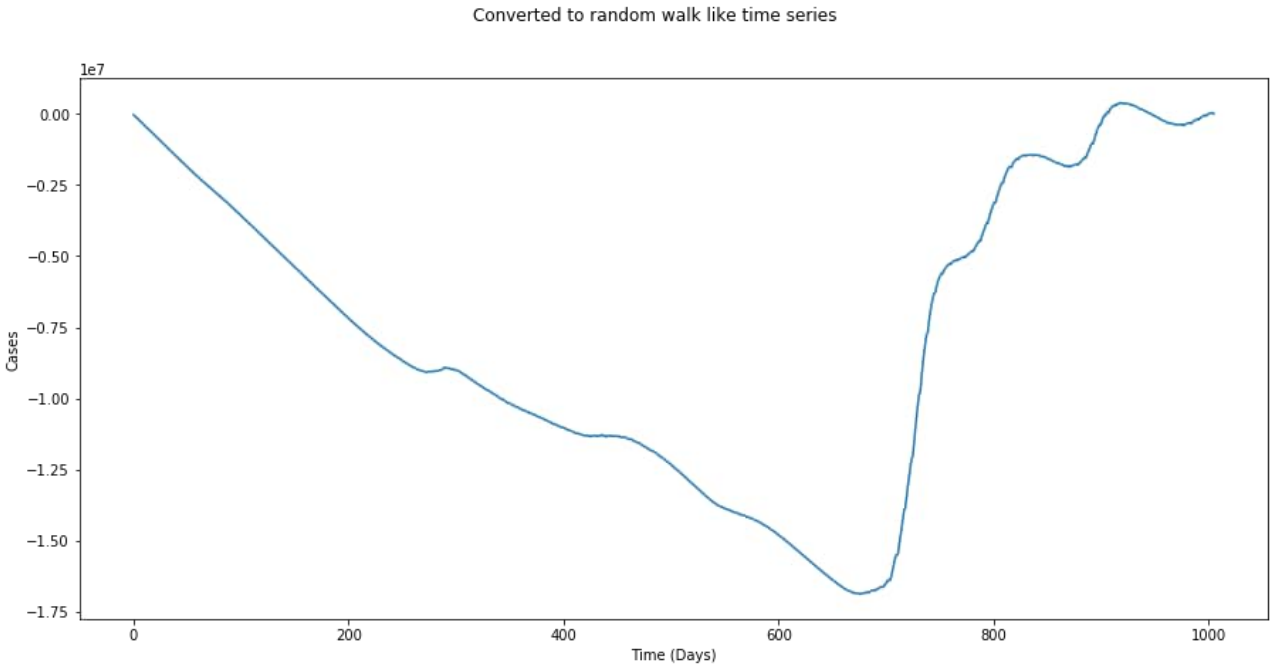

Рис. 4.2. Часовий ряд після перетворення з типу випадкового шуму до виду випадкового блукання.

Фрактальні характеристики випадкового блукання, вказаної на рис. 4.2. напряму відображають характеристики часового ряду на рис. 4.1. Спуски та підйоми у перетвореному часовому ряді вказують на періоди з низькою та високою захворюваністю на COVID-19.

Для аналізу варіації часового ряду зазвичай обчислюють варіацію як корінь від середнього значення для квадрату елементів:

$$RMS = \sqrt{avg\left(x^2(i)\right)}, \tag{11}$$

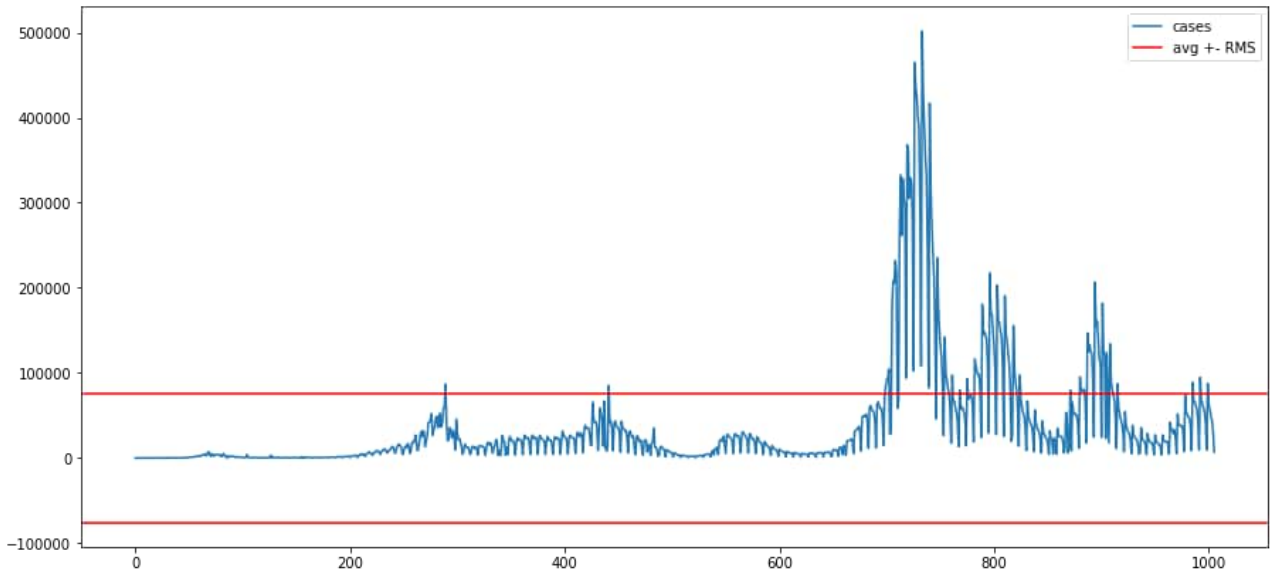

Рис. 4.3. Обчислення варіації часового ряду.

Потрібно зазначити, що RMS відображає тільки силу варіації, але не структуру. Мій часовий ряд містить області як з флуктуаціями великої сили, так і малої. Для того, щоб це врахувати його можна розділити на рівні частини однакового розміру, які не перетинаються і обчислити локальний RMS для кожної частини.

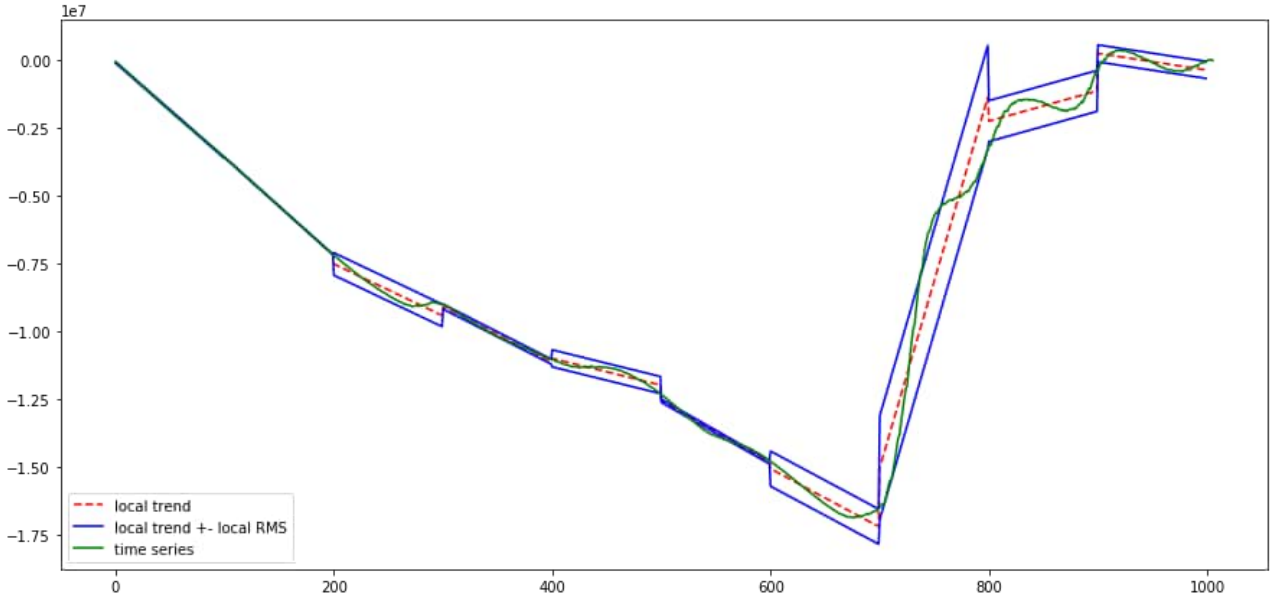

Рис.4.4. Обчислення локальної RMS часового ряду.

На рис. 4.4. часовий ряд було поділено на сегменти розміром у 100 елементів кожний. У кожного елементу було знайдено тренд і обчислено локальний RMS за такою формулою:

$$RMS = \sqrt{avg(Y - trend)^2}, \quad (12)$$

У даному випадку тренд обчислюється як пряма лінія, але також існують варіанти з квадратичним або кубічним трендом, де використовуються поліноми другої і третьої степені. Локальний RMS є основою для алгоритму DFA (detrended fluctuation analysis). У цьому алгоритмі вводиться функція, яка обчислює корінь середнього від квадратів локальних RMS:

$$F = \sqrt{avg(RMS^2)}, \quad (13)$$

Ця функція обчислюється для різного розміру сегментів. При цьому швидко зникаючі флуктуації будуть впливати на F при сегментах малого розміру, в той час як більші і повільніші – при сегментах великого розміру.

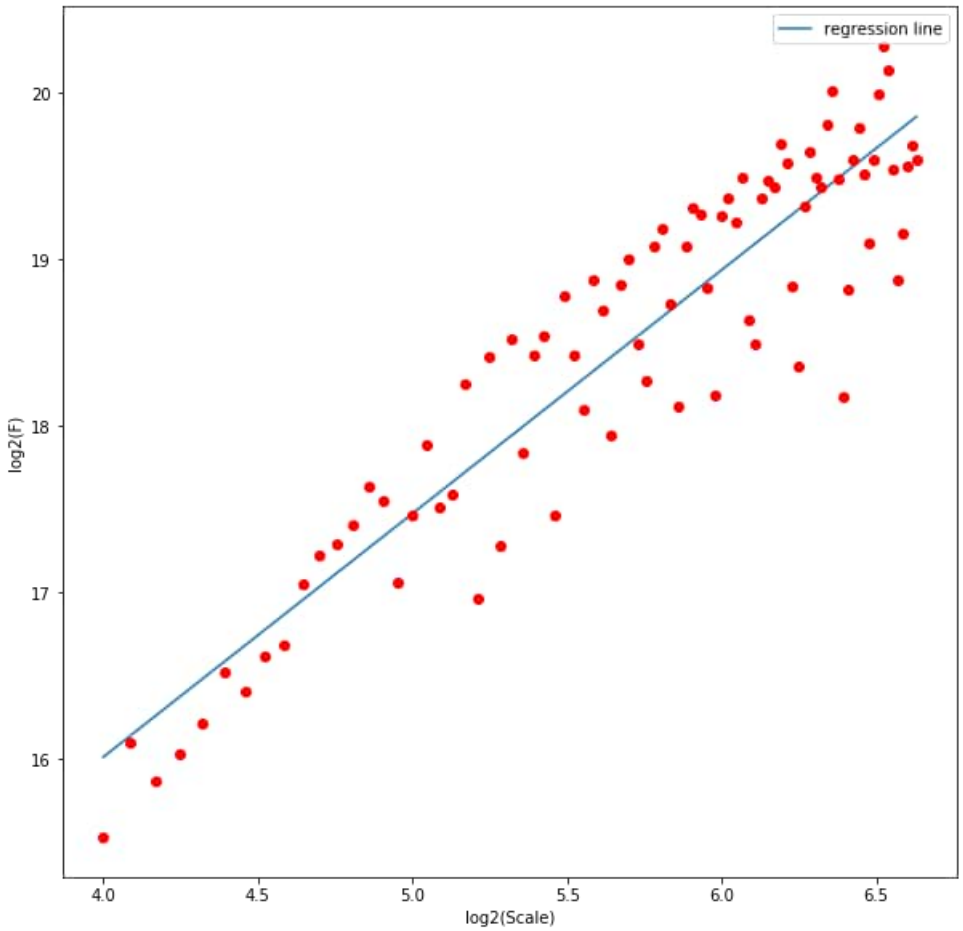

Рис. 4.5. Функція F та лінія регресії.

DFA обчислює монофрактальну структуру, як степенева залежність між F, обчислена при різному розмірі сегментів. Цю степеневу залежність

знаходять з нахилу лінії регресії, яка знаходиться шляхом обчислення лінійного поліному від $log_2(F)$ та $log_2(scale)$. Нахил лінії регресії називають показником Херста. Більше значення цього критерію означає довші та більш повільні варіації. У даному випадку показник Херста = 1.46. Потрібно зазначити, що монофрактальний та мультифрактальний часові ряди можуть мати похожі RMS та нахили ліній регресії, але їхні будови суттєво відрізняються. Мультифрактальний часовий ряд містить флуктуації різної величини, тоді як монофрактальний – ні. Код алгоритму MFA і MFDFA, розбитий на етапи, показаний у додатках А, Б і В. У алгоритмі MFDFA вводиться порядок статистичного моменту q, який змінює формулу F до такого виду:

$$F = \sqrt[q]{avg(RMS^q)}, \quad (14)$$

Відсутність флуктуації надзвичайно великих та малих величин призводить до рівномірного розподілу монофрактального часового ряду, який можна описати одним лиш порядком статистичного моменту $q = 2$.

У моєму випадку було вибрано дані параметри:

scale = [10, 100],

q = [-5, 5],

m = 1,

Це означає що часовий ряд буде ділитись на сегменти розміром від 10 до 100 елементів, тренд даних сегментів буде обчислюватись поліномами першого степеню, а порядок статистичного моменту буде в межах від -5 до 5. При даних параметрах я отримав такі результати:

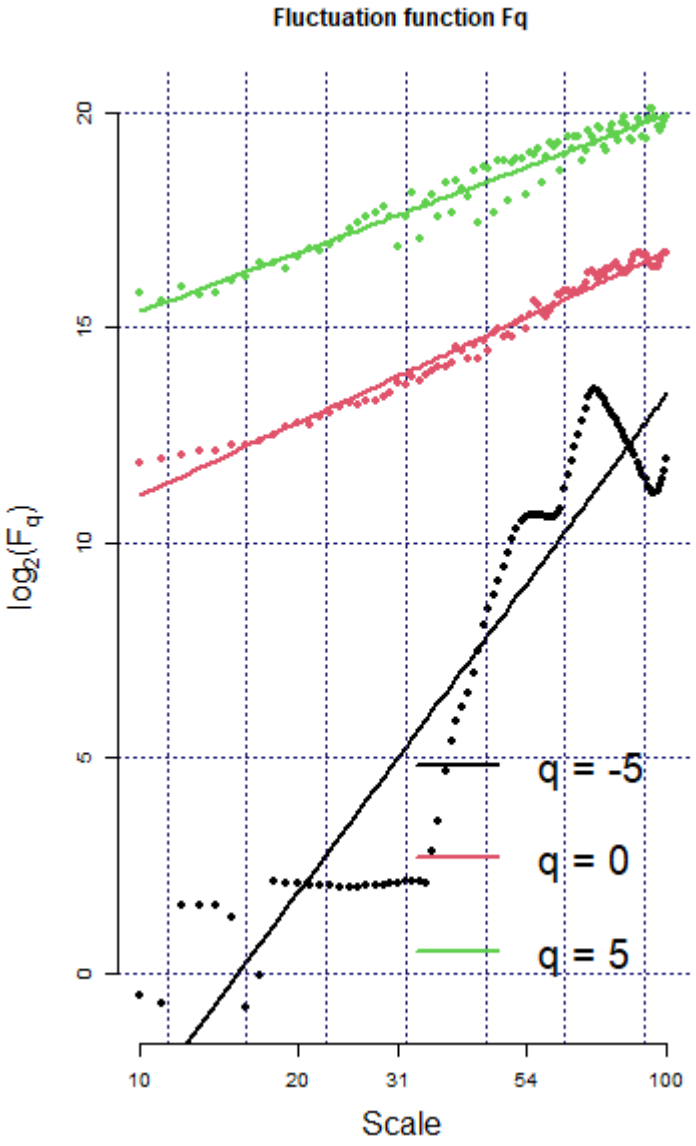

Рис.4.6 Функція флуктуацій та лінії регресії при різних значеннях q.

На рис. 4.6. Показано зміну куту нахилу лінії регресії при різних значеннях q. Залежність куту нахилу від q вказує на те, що часовий ряд є мультифрактальним. Різниця між F_q при додатніх та від'ємних q найбільша при малих розмірах сегментів. Це пояснюється тим, що малі сегменти здатні розрізняти локальні періоди з великими і малими флуктуаціями, так як вони можуть поміститися всередині цих періодів. Толі як великі сегменти можуть перекрити декілька періодів з великими і малими флуктуаціями та усереднити їх. У випадку з монофрактальними часовими рядами, вони не мають періодів з

великими і малими флуктуаціями. Тому їх лінії регресії не залежать від q і знаходяться паралельно одна одній.

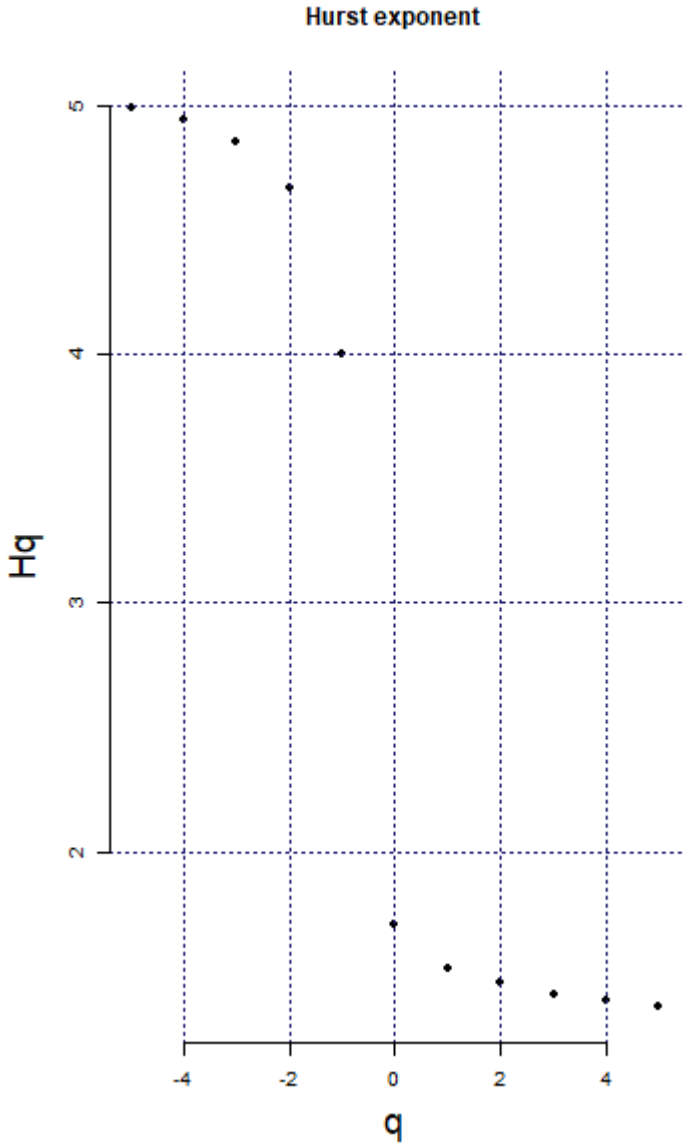

Рис.4.7. Спектр фрактальних розмірностей.

На рис.4.7. відображаються показники Херста порядку q при різних значеннях q. Даний варіант відображення дозволяє більш наглядно подивитись на зміну куту нахилу ліній регресії. З попереднього графіку можна визначити, що для монофрактальних часових рядів показник Херста буде константою при всіх значеннях q. Також варто зазначити, що у нашому випадку при великих значеннях q графік h_q також стає подібним до константи. Загальне зниження показників Херста вказує на те, що сегменти з малими флуктуаціями містять

структуру подібну до випадкового блукання, тоді як сегменти з великими флуктуаціями – структуру подібну до випадкового шуму.

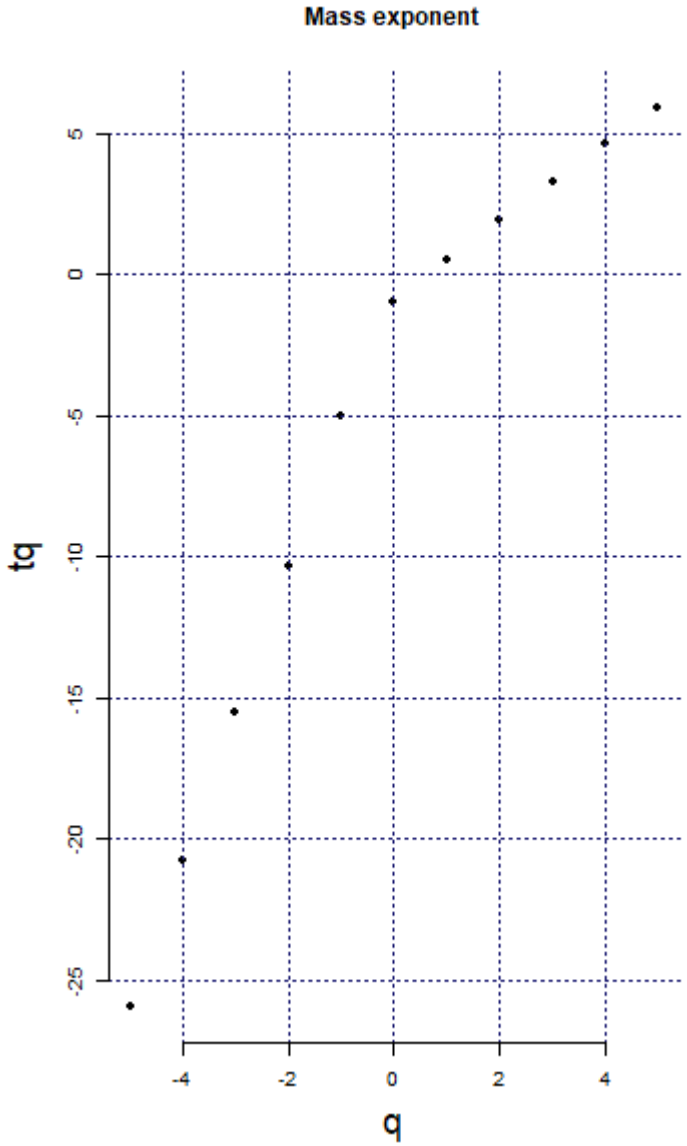

Рис. 4.9. Спектр скейлінгових експонент.

Окрім показника Херста порядку q існують також інші методи параметризації мультифрактальної структури часового ряду. Спектр скейлінгових експонент обчислюється за формулою:

$$\tau_q = qH_q - 1, \quad (15)$$

Він використовується для обчислення показника сингулярності порядку q та виміру сингулярності порядку q за такими формулами:

$$h_q = \frac{d\tau_q}{dq}, \tag{16}$$

$$D_q = qh_q - \tau_q, \tag{17}$$

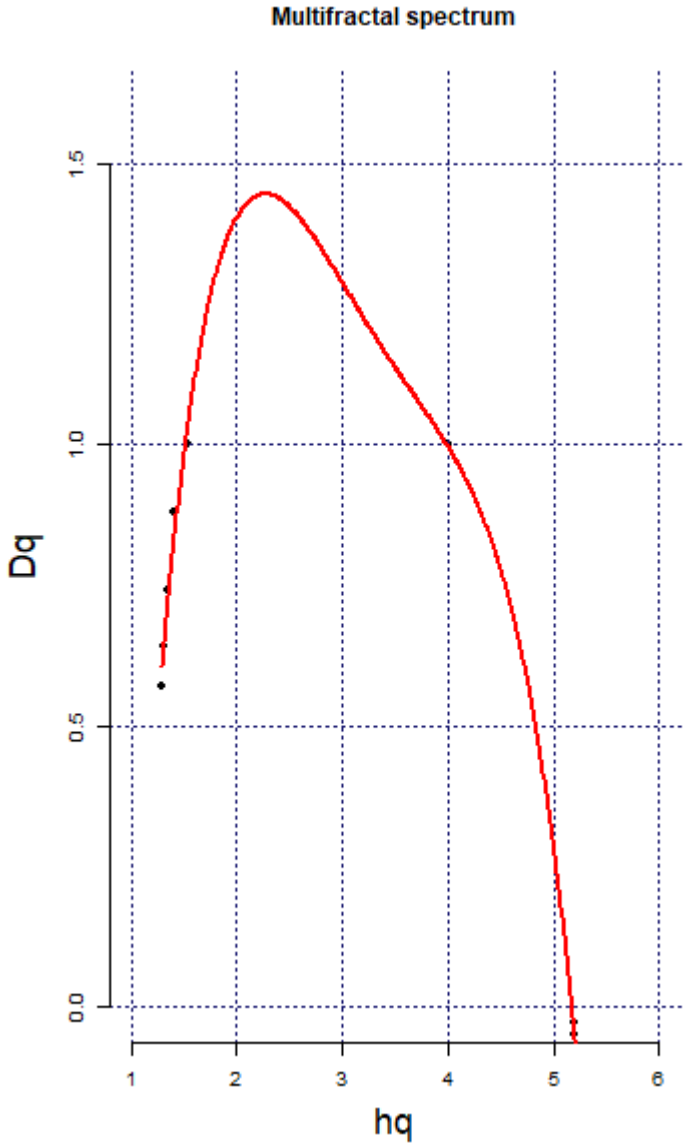

Рис. 4.10 Спектр сингулярностей

Монофрактальні часові ряди мають скейлінговий критерій τ_q з лінійною залежністю від q. Через це спектр сингулярностей у них буде зведений до точки, або до дуги дуже малого розміру. В той час як мультифрактальні часові ряди мають скейлінговий критерій τ_q з залежністю від q у вигляді кривої лінії (рис.

4.9.) та показник сингулярності h_q, який поступово зменшується. В результаті цього спектр сингулярностей такого ряду буде у вигляді великої дуги, ширина якої є різниця між найбільшим та найменшим h_q. З рис.4.10 видно, що дуга спектру сингулярностей не є симетричною. Різна довжина хвостів дуги вказує на низьку чутливість часового ряду до флуктуацій великої або малої величини. У даному випадку дуга має довший правий хвіст, тому часовий ряд є нечутливим до великих флуктуацій.

4.2. Програмна реалізація моделі ARFIMA.

Бібліотека arfima призначена для симуляції, побудови та прогнозування часових рядів на одноіменній моделі. Містить широкий діапазон команд, які полегшують роботу з даними, які нам потрібно обчислити і вивести.

Функція приймає вектор *z* – часовий ряд з одною змінною, *p* – AR параметр, *q* – MA параметр та *d* – I параметр. Для часового ряду з новими випадками захворювань в Україні було обрано такі значення:

p – 8

q – 0.604

d – 3

Також для створення прогнозу було використано функцію predict, яка приймає object – модель і horizon – кількість спрогнозованих елементів. У моєму випадку horizon = 14, тобто моя система буде прогнозувати результати на наступні два тижні.

4.3. Результати досліджень.

Спрогнозовані дані будуть візуально представлені у вигляді графіків, де чорним кольором позначено дані з реального датасету, а червоним — спрогнозовані дані. Також буде виведено середнє квадратичне відхилення (RMSE) та середнє абсолютне відсоткове відхилення (MAPE) між спрогнозованими та реальними значеннями. На рис 4.11. та таблиці 4.1.

показано результати прогнозів з використанням моделі ARFIMA з визначеними раніше показниками.

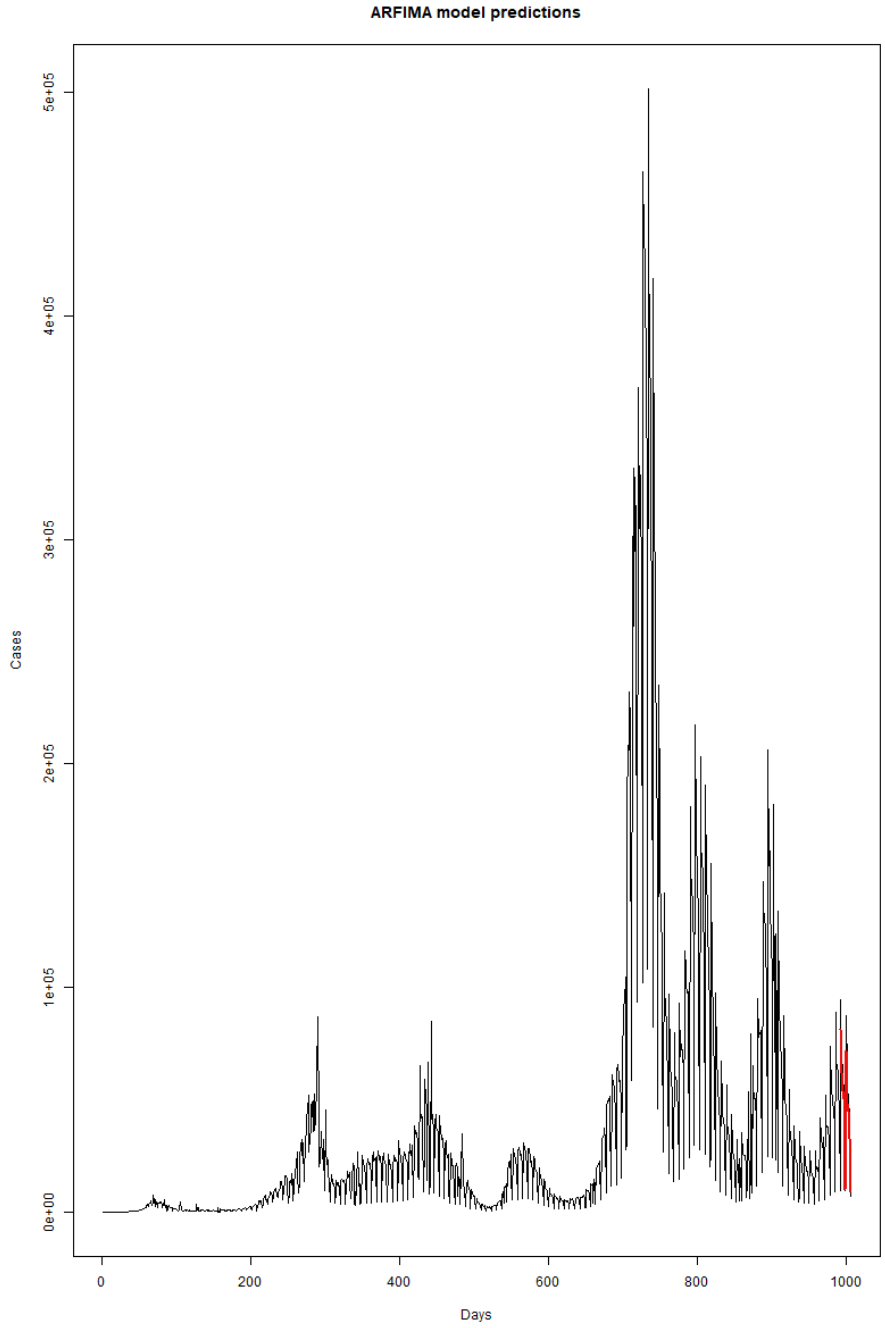

Рис. 4.11. Спрогнозовані результати нових випадків захворювань, виконані моделлю ARFIMA.

Таблиця 4.1. Похибки при прогнозуванні моделлю ARFIMA

RMSE	MAPE
7101.96	10.64

4.4. Аналіз існуючих результатів.

Для порівняння було використано такі моделі:

- VAR
- VARMA
- ARIMA
- ARFIMA з автоматично визначеними параметрами

Порівняння проходило з використанням старого датасету захворюваності в Україні, який містить 500 записів і на даний момент втратив актуальність. Датасет взятий з ресурсу Kaggle [10]. Також при порівнянні використовувались як часові ряди з показниками захворюваності так і з показниками смертності. Кількість тестових даних була 30, тренувальних – 470. На рис. 4.12., рис. 4.13. та таблиці 4.2. показані результати прогнозування моделлю ARFIMA з параметрами, підібраними вручну для часових рядів нових випадків захворювання і смертності від COVID-19.

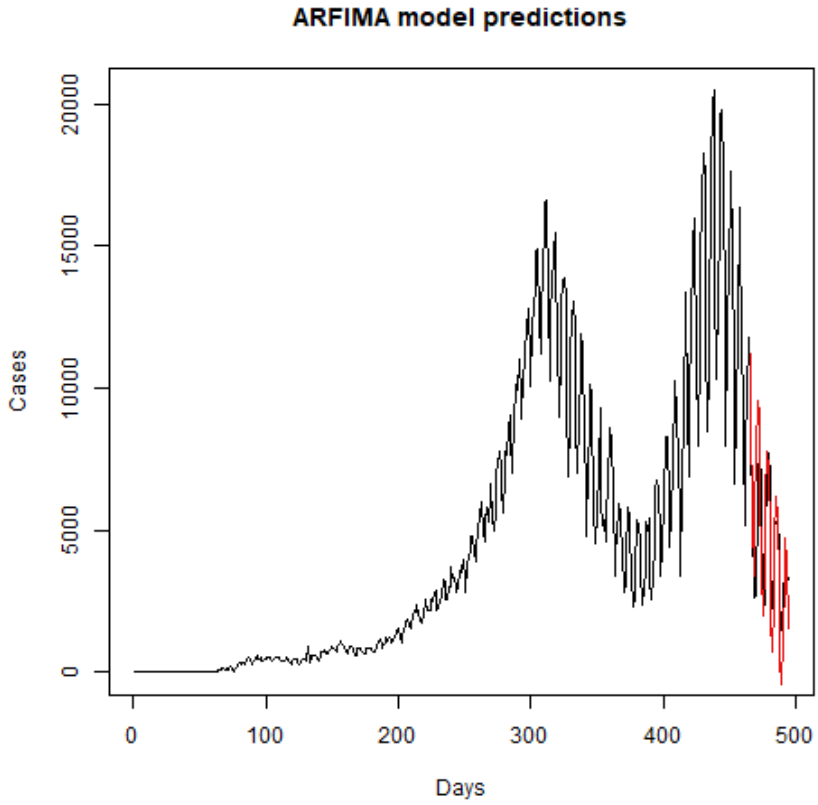

Рис. 4.12. Спрогнозовані результати нових випадків захворювань, виконані моделлю ARFIMA.

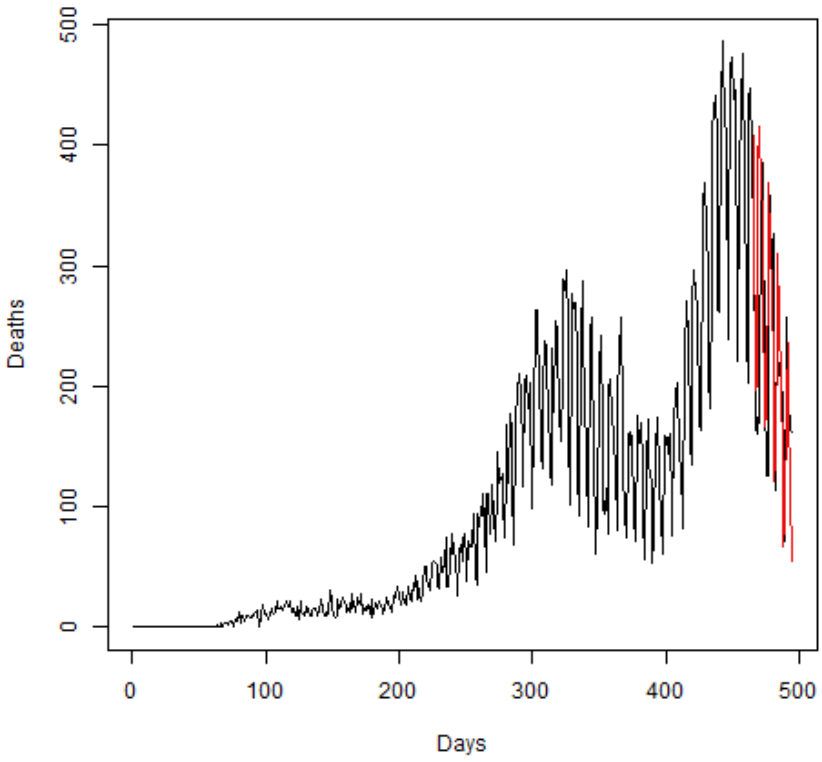

Рис. 4.13. Спрогнозовані результати нових смертей в Україні, виконані моделлю ARFIMA.

Таблиця 4.2. Похибки при прогнозуванні моделлю ARFIMA

Dataset	RMSE	MAPE
Cases	1783.01	41.42
Deaths	84.34	31.77

На рис. 4.14. та таблиці 4.3 показано результати прогнозів з використанням моделі VAR з автоматично підібраними параметрами.

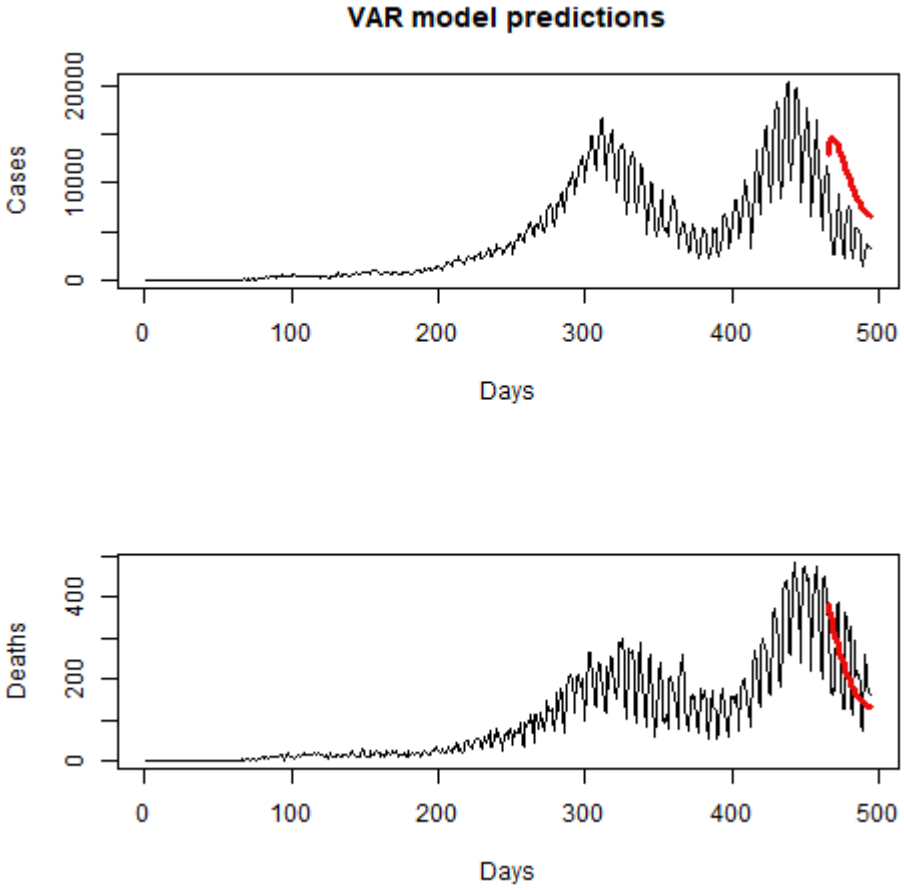

Рис. 4.14. Спрогнозовані результати нових випадків та смертей виконані моделлю VAR.

Таблиця 4.3. Похибки при прогнозуванні моделлю VAR

Dataset	RMSE	MAPE
Cases	6485.08	166.45
Deaths	100.46	46.46

На рис. 4.15. та таблиці 4.4 показано результати прогнозів з використанням моделі VARMA з автоматично підібраними параметрами.

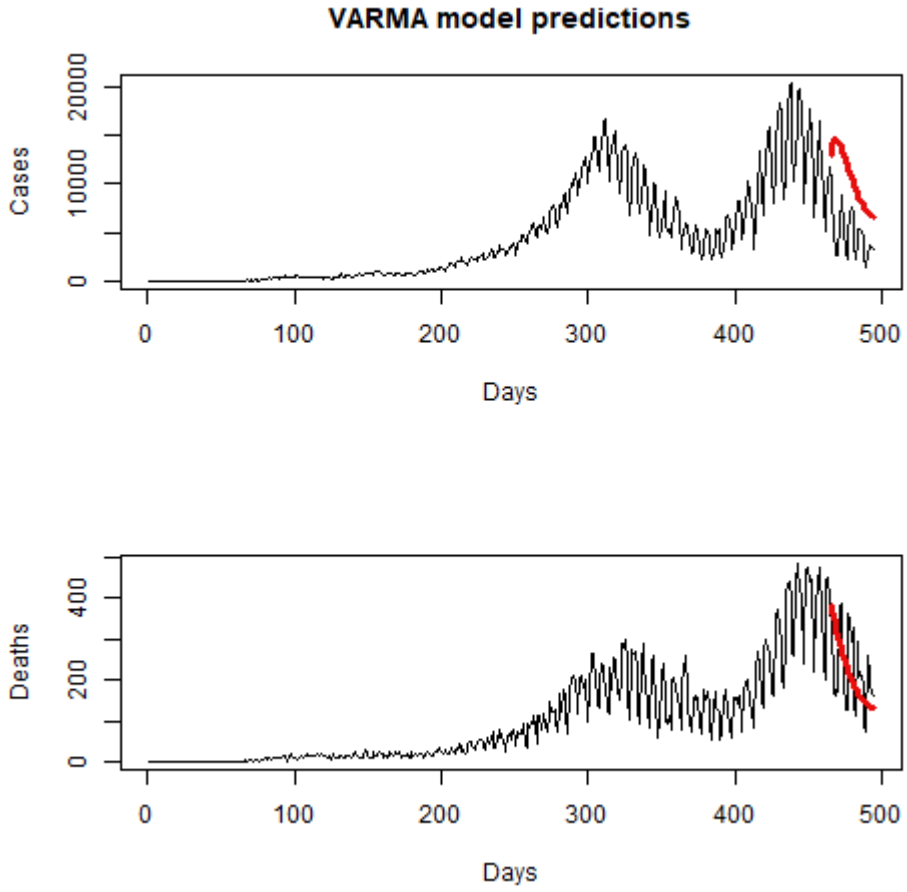

Рис. 4.15. Спрогнозовані результати нових випадків та смертей виконані моделлю VARMA.

Таблиця 4.4. Похибки при прогнозуванні моделлю VARMA

Dataset	RMSE	MAPE
Cases	6461.83	165.84
Deaths	100.51	46.37

На рис. 4.16., рис. 4.17. та таблиці 4.5 показано результати прогнозів з використанням моделі ARIMA з автоматично підібраними параметрами.

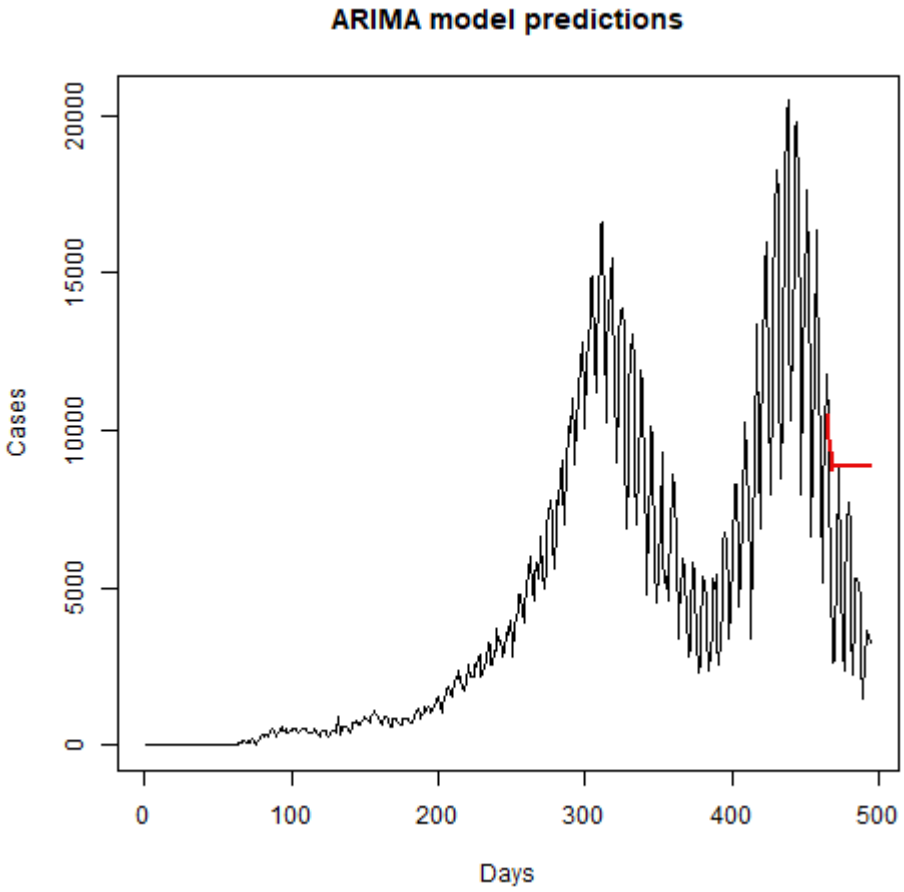

Рис. 4.16. Спрогнозовані результати нових випадків захворювань, виконані моделлю ARIMA.

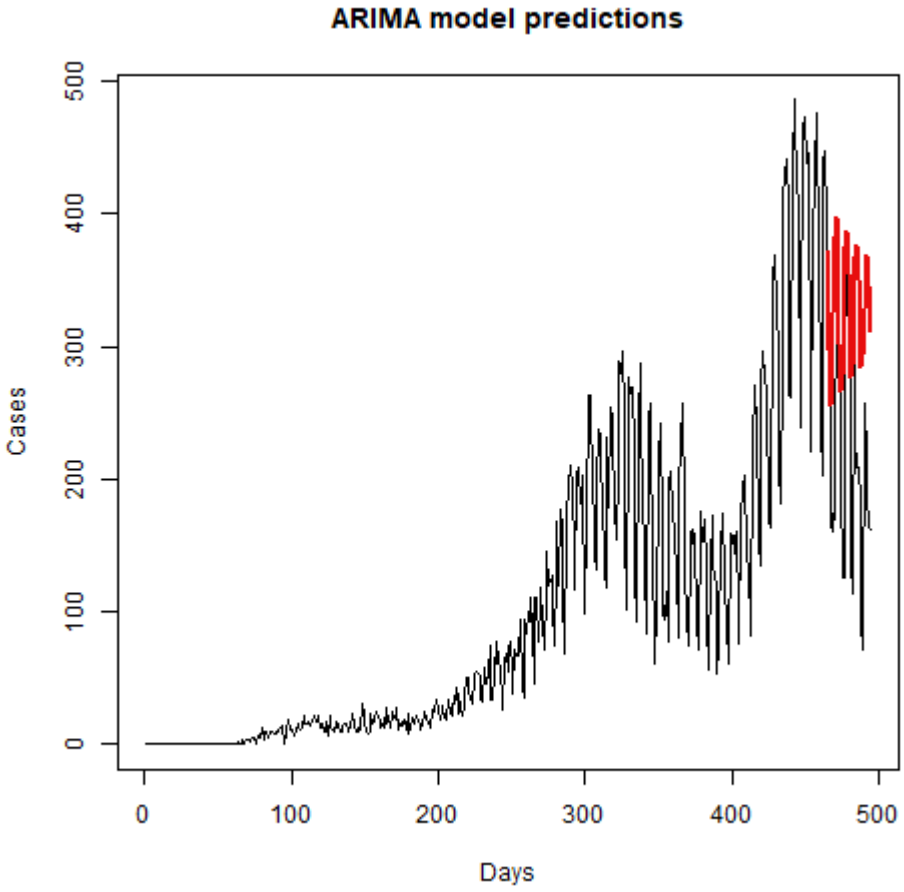

Рис. 4.17. Спрогнозовані результати нових смертей в Україні, виконані моделлю ARIMA.

Таблиця 4.5. Похибки при прогнозуванні моделлю ARIMA

Dataset	RMSE	MAPE
Cases	4646.75	131.52
Deaths	130.04	73.40

На рис. 4.18., рис. 4.19. та таблиці 4.6. показано результати прогнозів з використанням моделі ARFIMA з автоматично визначеними параметрами.

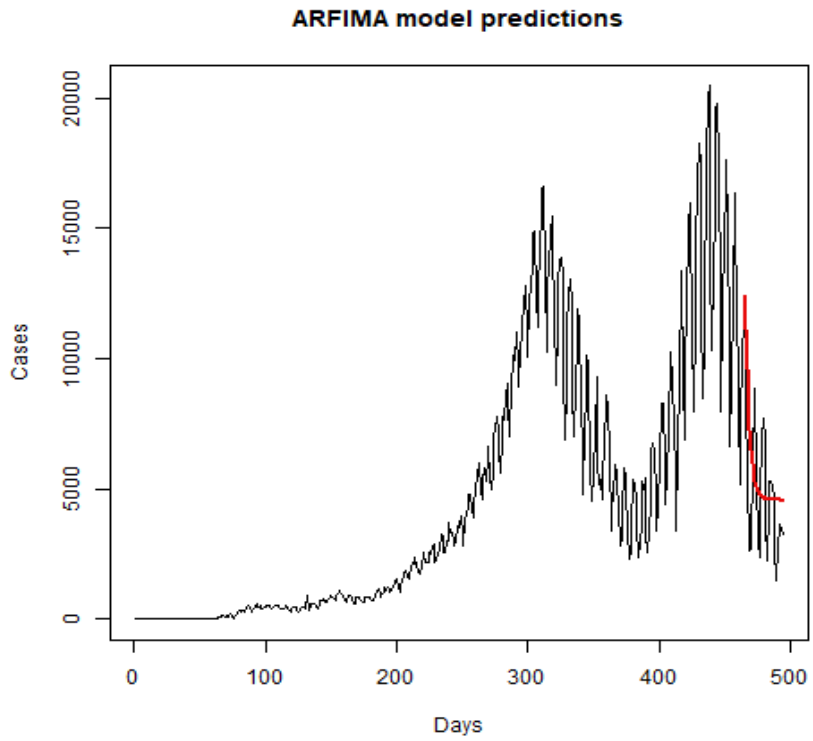

Рис. 4.18. Спрогнозовані результати нових випадків захворювань, виконані моделлю ARFIMA (auto).

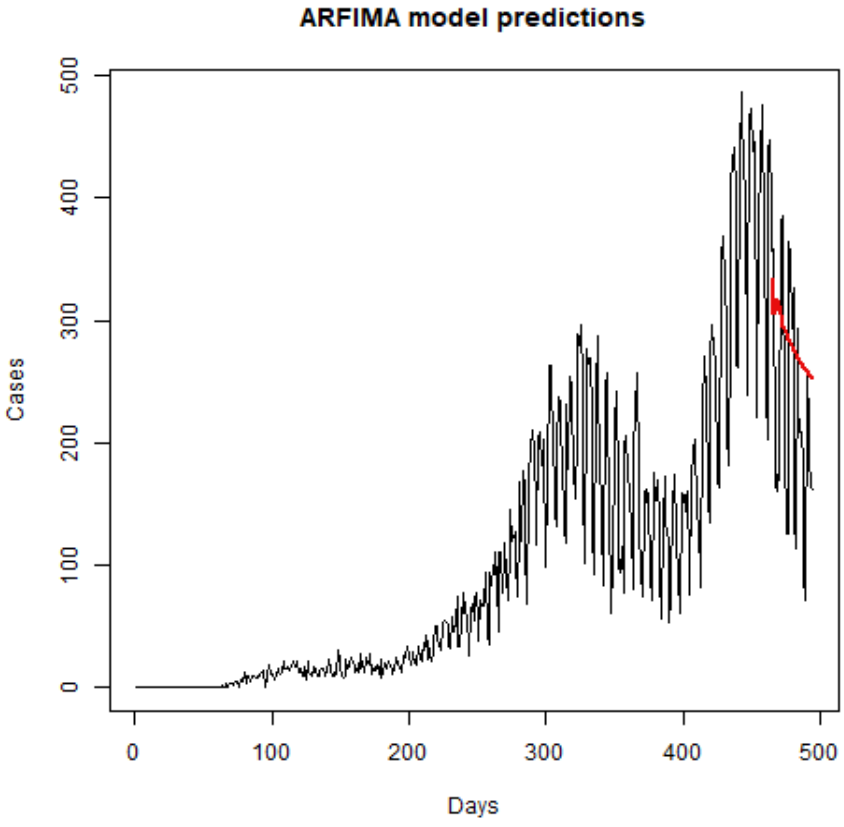

Рис. 4.19. Спрогнозовані результати нових смертей в Україні, виконані моделлю ARFIMA (auto).

Таблиця 4.6. Похибки при прогнозуванні моделлю ARFIMA(auto)

Dataset	RMSE	MAPE
Cases	2421.39	55.52
Deaths	104.86	59.11

Параметри інших моделей були вибрані автоматично на основі алгоритмів, які використовуються в бібліотеках, де ці моделі знаходяться. На основі результатів можна побачити, що моя модель змогла скласти прогнози з меншими похибками відносно інших моделей, а також з цією ж моделлю, але автоматично визначеними параметрами.

4.5. Висновки.

Мультифрактальний аналіз за допомогою алгоритму MFDFA показав, що даний часовий ряд має мультифрактальну структуру. Розподіл показників Херста порядку q вказує на те, що у часовому ряді флуктуації малої величини містять фрактальну структуру, подібну до випадкового блукання, а флуктуації великої величини – структуру подібну до випадкового шуму. Також нерівномірність мультифрактального спектру вказує на нечутливість часового ряду до великих флуктуацій.

При складанні прогнозів поширення COVID-19 на 14 днів, модель ARFIMA мала помилку по кореневому середньоквадратичному відхиленню 7102, та по середній абсолютній відсотковій помилці 10.6%. При порівнянні з іншими моделями прогнозування модель ARFIMA показала найкращий результат.

ВИСНОВКИ

Під час виконання даної роботи було проаналізовано динаміку поширення вірусу. Серед наявних наборів даних було обрано дані по захворюваності в Європі. Серед країн Європи були вибрано Францію та створено датасет, який налічує близько 1000 елементів, починаючи з січня 2020 року, закінчуючи жовтнем 2022 року. Також проведено чистку ряду від аномалій і неправильних даних. Для цього було використано мову програмування Python.

Для визначення наявності довгої пам’яті було складено графіки автокореляції та часткової автокореляції. При аналізі цих графіків було визначено, що мій датасет має властивості довгострокової залежності.

Використавши фрактальну модель ARFIMA було складено прогноз поширення хвороби на наступні 14 днів. Також було проведено порівняння результатів прогнозування з іншими моделями. Основними критеріями порівняння були використані середнє квадратичне відхилення RMSE та абсолютне відсоткове відхилення MAPE між спрогнозованими і реальними значеннями. Результати порівняння показали, що модель ARFIMA склала прогноз з найменшими похибками і відхиленнями. Для програмної реалізації були використано мову програмування R та бібліотеку arfima.

Було проведено мультифрактальний аналіз датасету. Для цього було використано алгоритм MFDFA. Результати вказують на те, що часовий ряд має мультифрактальну структуру. При цьому флуктуації малої величини містять структуру подібну до випадкового блукання, а флуктуації великої величини – структуру подібну до випадкового шуму. В загальному датасет є нечутливим до великих флуктуацій. Для програмної реалізації було використано мову програмування R та бібліотеку MFDFA.

ДОДАТКИ

Додаток А. Реалізація знаходження локального RMS

```
X = []
X_avg = mean(cases)

for i in range(len(cases)):
    X.append(cases[i] - X_avg)

for i in range(1, len(X)):
    X[i] += X[i-1]

out = plt.figure()
out.set_figwidth(15)
out.set_figheight(7)
plt.ylabel('Cases')
plt.xlabel('Time (Days)')
plt.plot(X)
plt.suptitle("Converted to random walk like time series")
plt.show()
```

Перетворення часового ряду cases який є подібним до випадкового шуму до часового ряду X який є подібним на випадкове блукання.

```
def coefs_to_Y(coefs, Xs):
    #returns a polynomial for X values for coefs provided
    result = []
    for i in range(len(Xs)):
        y = Xs[i]*coefs[1] + coefs[0]
        result.append(y)
    return result

scale = 100
m = 1
segments = int(len(X)/scale)
Idx_i = []
for i in range(1,segments+1):
    Idx_start=(i-1)*scale
    Idx_stop=i*scale
    Idx_i.append([list(range(Idx_start,Idx_stop)),X[Idx_start:Idx_stop]])

res = []

for i in range(segments):
    C = np.polynomial.polynomial.Polynomial.fit(Idx_i[i][0], Idx_i[i][1], m)
    res.append(C.convert().coef)

local_RMS = []
lines = []
lines_out = []
for i in range(len(res)):
    seg = coefs_to_Y(res[i],Idx_i[i][0])
    for j in range(scale):
        tmp_sqr = []
        tmp_sqr.append(pow((Idx_i[i][1][j]-seg[j]),2))
    local_RMS.append(pow(mean(tmp_sqr),0.5))
    lines.append(seg)
    lines_out.extend(seg)
```

Обчислення локального RMS при сегментах розміру 100 та використання поліномів першого степеню для обчислення трендів.

Додаток Б. Реалізація алгоритму DFA.

```
def scale_RMS(X, scale, m):
    segments = int(len(X)/scale)
    Idx_i = []
    for i in range(1,segments+1):
        Idx_start=(i-1)*scale
        Idx_stop=i*scale
        Idx_i.append([list(range(Idx_start,Idx_stop)),X[Idx_start:Idx_stop]])
    res = []
    for i in range(segments):
        C = np.polynomial.polynomial.Polynomial.fit(Idx_i[i][0], Idx_i[i][1], m)
        res.append(C.convert().coef)
    local_RMS = []
    for i in range(len(res)):
        seg = coefs_to_Y(res[i],Idx_i[i][0])
        for j in range(scale):
            tmp_sqr = []
            tmp_sqr.append(pow((Idx_i[i][1][j]-seg[j]),2))
        local_RMS.append(pow(mean(tmp_sqr),0.5))
    return local_RMS

scale = list(range(16,100))
m = 1
Fi = []
for i in range(len(scale)):
    rms = scale_RMS(X, scale[i], m)
    for j in range(len(rms)):
        rms[j] = pow(rms[j], 2)
    Fi.append(pow(mean(rms),0.5))

scale_log2 = []
F_log2 = []
for i in range(len(scale)):
    scale_log2.append(math.log2(scale[i]))
    F_log2.append(math.log2(Fi[i]))

C = np.polynomial.polynomial.Polynomial.fit(scale_log2, F_log2, 1)
H = C.convert().coef[1]

out = plt.figure()
out.set_figwidth(10)
out.set_figheight(10)
lines = coefs_to_Y(C.convert().coef,scale_log2)
plt.plot(scale_log2,lines, label = "regression line")
plt.plot(scale_log2,F_log2, 'ro')
plt.ylabel('log2(F)')
plt.xlabel('log2(Scale)')
leg = plt.legend()
plt.show()
```

Обчислення функції флуктуацій F_q яка позначена як Fi та показника Херста, який позначений як H. Метод scale_RMS, обчислює локальний RMS при заданому параметрі Scale.

Додаток В. Реалізація алгоритму MFDFA.

```
def MFDFA(X, scale, q, m):
    Fqs = []
    for ns in range(len(scale)):
        rms = scale_RMS(X, scale[ns], m)
        Fq = []
        QRMS = []
        for nq in range (len(q)):
            qRMS = []
            for i in range(len(rms)):
                qRMS.append(pow(rms[i], q[nq]))

            QRMS.append(qRMS)
            if q[nq] == 0:
                tmp = []
                for i in range(len(rms)):
                    tmp.append(math.log(pow(rms[i], 2)))
                Fq.append(math.exp(0.5 * mean(tmp)))
            else:
                Fq.append(pow(mean(qRMS), 1/q[nq]))
        Fqs.append(Fq)
    return Fqs

scale = list(range(10,100))
q=list(range(-5,6))
m = 1
Fq = MFDFA(X, scale, q, m)
```

Метод MFDFA приймає параметри scale – розміри сегментів, q – порядок статистичного моменту і m – степінь поліномів, за яких обчислюється тренд, та повертає функції флуктуацій при заданих параметрах.

```
Hq = []
Hq_lines = []

scale_log2 = []
for i in range(len(scale)):
    scale_log2.append(math.log2(scale[i]))

for qn in range(len(q)):
    F_log2 = []
    for i in range(len(scale)):
        F_log2.append(math.log2(Fqs[qn][i]))
    C = np.polynomial.polynomial.Polynomial.fit(scale_log2, F_log2, 1)
    H = C.convert().coef[1]
    Hq.append(H)
    lines = coefs_to_Y(C.convert().coef,scale_log2)
    Hq_lines.append(lines)

plt.plot(q,Hq,'o')
plt.show()
```

Обчислення ліній регресії та показників Херста порядку q.

```
tq = []
for i in range(len(q)):
    tq.append(Hq[i]*q[i]-1)

out = plt.figure()
plt.plot(q,tq,'o')
plt.show()
```

Обчислення спектру скейлінгових експонент.

```
hq = []
for i in range(1, len(tq)):
    diff_tq = tq[i] - tq[i-1]
    diff_q = q[i] - q[i-1]
    hq.append(diff_tq/diff_q)

Dq = []
for i in range(len(hq)):
    Dq.append((q[i]*hq[i])-tq[i])

plt.plot(hq,Dq,'o')
plt.show()
```

Обчислення мультифрактального спектру.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

[1] WHO-convened Global Study of the Origins of SARS-CoV-2 [Електронний ресурс]/ World Health Organization, https://apo.org.au/node/311637

[2] А. Н. Павлов, В. С. Анищенко, Мультифрактальный анализ сложных сигналов, УФН, 2007, том 177, номер 8, 859–876, https://doi.org/10.3367/UFNr.0177.200708d.0859

[3] Liu, K.; Chen, Y.; Zhang, X. An Evaluation of ARFIMA (Autoregressive Fractional Integral Moving Average) Programs. *Axioms* **2017**, *6*, 16. https://doi.org/10.3390/axioms6020016

[4] Jan W. Kantelhardt, Stephan A. Zschiegner, Eva Koscielny-Bunde, Shlomo Havlin, Armin Bunde, H.Eugene Stanley, Multifractal detrended fluctuation analysis of nonstationary time series, Physica A: Statistical Mechanics and its Applications, Volume 316, Issues 1–4, 2002, Pages 87-114, https://doi.org/10.1016/S0378-4371(02)01383-3.

[5] Veenstra, Justin Quinn, "Persistence and Anti-persistence: Theory and Software" (2013). *Electronic Thesis and Dissertation Repository*. 1119. https://ir.lib.uwo.ca/etd/1119

[6] Kantelhardt, J.W. (2012). Fractal and Multifractal Time Series. In: Meyers, R. (eds) Mathematics of Complexity and Dynamical Systems. Springer, New York, NY. https://doi.org/10.1007/978-1-4614-1806-1_30

[7] Ihlen Espen, Introduction to multifractal detrended fluctuation analysis in Matlab, Frontiers in Physiology, Volume 3, 2012, https://doi.org/10.3389/fphys.2012.00141

[8] Historical data on the daily number of new reported COVID-19 cases and deaths worldwide [Електронний ресурс] / https://www.ecdc.europa.eu/en/publications-data/download-todays-data-geographic-distribution-covid-19-cases-worldwide

[9] Laib, Mohamed & Telesca, Luciano & Kanevski, Mikhail. (2017). MFDFA: MultiFractal Detrended Fluctuation Analysis. https://doi.org/10.13140/RG.2.2.13692.23686.

[10] Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Inf Dis. 20(5):533-534. https://doi.org/10.1016/S1473-3099(20)30120-1