

В.В. Скалозуб, В.М. Горячкін, І.В. Клименко, Д.О. Шаповал

**МОДЕЛІ ТА ПРОЦЕДУРИ КЛАСИФІКАЦІЇ І ПРОГНОЗУВАННЯ
НЕДЕТЕРМІНОВАНИХ ПРОЦЕСІВ
ЗА ПОКАЗНИКАМИ ХАОТИЧНОЇ ДИНАМІКИ**

Анотація. У статті досліджуються процеси класифікації, моделювання та короткострокового прогнозування недетермінованих часових послідовностей, які представлені антиперсистентними часовими рядами (АЧР). Предмет аналізу – процедури класифікації та прогнозування параметрів таких моделей. Метою роботи є підвищення ефективності та точності методів і алгоритмів класифікації, моделювання та прогнозування АЧР. За допомогою агрегування рівнів АЧР розроблено коректні математичні моделі класифікації недетермінованих часових послідовностей, а також сформовано алгоритмічні та програмні засоби їх реалізації. Проведено порівняльний аналіз чисельної ефективності алгоритмів класифікації АЧР, реалізовано завдання короткострокового прогнозування АЧР. У статті також представлено інструментальне програмне середовище, що забезпечує коректне вивчення алгоритмів моделювання та класифікації антиперсистентних часових рядів (АЧР).

Ключові слова: антиперсистентні часові ряди, класифікація ЧР, моделювання, інтерполяція, короткострокове прогнозування, точність алгоритмів класифікації, програмне забезпечення.

Вступ та постановка проблеми. Всебічний аналіз станів та ефективне управління складними недетермінованими процесами, зокрема залізничного транспорту (ЗТ), з використанням лише класичних методів математико-статистичного моделювання є неможливими. Розподіленість та складність процесів ЗТ (поїздоутворення, технології здійснення процесів вантажних перевезень а також ремонту, функціонування транспортної інфраструктури та її об'єктів ін.) призводять до виникнення подій і ситуацій з принципово недетермінованими параметрами. Перелічені та інші фактори обумовлюють причини відсутності загальних математичних моделей для багатьох процесів ЗТ технолого-економічного змісту. У зв'язку зі складністю і недетермінованістю численних процесів ЗТ одним із основних способів їх первинного опису є представлення даних у вигляді упорядкованих за етапами та інтервалами значень основних вимірюваних характеристик часових рядів (ЧР). Прикладами таких ЧР для залізниці можуть бути параметри вагонопотоків різних

категорій на полігонах і станціях, оцінки витрат різних видів ресурсів за періодами, параметри вантажопотоків, послідовності подій процесів експлуатації та ін.

Моделі і методи аналізу та прогнозування станів і параметрів складних систем на основі ЧР отримали визнання і широке поширення, знаходять нові області застосування і форми реалізації. Відзначається що ЧР процесів ЗТ і вимоги до результатів їх аналізу мають специфіку у порівнянні з аналізом ЧР природних процесів (біологічні та атмосферні явища, кардіологічні процеси тощо). Основні з них – по-перше, що ЧР є «короткими» (до 250 рівнів); по-друге, вимога результативності щодо «придатності» процедур аналізу для реалізації систематичного та достовірного виробничо-технічного прогнозування і планування.

Для дослідження загальних властивостей ЧР натеper використовуються моделі та методи фрактального аналізу, зокрема, обчислення показника Херста. (H), за значеннями якого встановлюються категорії процесів ЧР (персистентні або трендостійкі, антиперсистентні, де очікуються «злами» тенденції, процеси випадкового характеру). Для дослідження антиперсистентних ЧР (АЧР) були розроблені спеціалізовані процедури, які за рахунок узагальнення послідовних рівнів АЧР дозволили диференціювати та класифікувати такі недетерміновані процеси.

Формування та інтерпретація послідовностей даних АЧР потребує визначення можливості коректного формування комп'ютерних моделей, процедур оцінки параметрів станів процесів, методики дослідження властивостей ЧР, точності та достовірності побудованих моделей тощо. Розвиток досліджень щодо моделей, алгоритмів і програмних засобів, призначених для аналізу, класифікації та прогнозування АЧР натеper має науковий і практичний інтерес, що свідчить про актуальність представленої роботи. В статті вирішуються питання що забезпечують розробку алгоритмічних та інструментальних програмних засобів, а також виконання порівняльного аналізу чисельної ефективності алгоритмів коректного дослідження і класифікації АЧР, формування процедур інтерполяції та короткострокового прогнозування АЧР.

Аналіз останніх досліджень і публікацій. Дослідженням властивостей складних процесів та систем, представлених часовими рядами (ЧР), а також методам та процедурам які для цього застосовуються, постійно приділяється увага, зокрема в роботах [1 – 6,13,14]. При цьому виділяється широкий спектр напрямків щодо цілей та інструментарію. В [1] відзначено що вимірювання складності та передбачуваності ЧР іноді є наріжним каменем для пропозицій

щодо розв'язання проблем. Як метод вимірювання складності та передбачуваності ЧР у багатьох областях широко застосовується багатомасштабна ентропія (MSE), У роботі [1] розроблено нову функцію подібності – гнучку багатомасштабну ентропію (FMSE), призначену для вимірювань схожості двох підпоследовностей. Пропонований метод FMSE здатний підвищити продуктивність процедур аналізу часових рядів.

У різних галузях досліджень ентропія була поширеним індексом для кількісної оцінки складності ЧР. В [2] вводиться нова форма – ентропія збільшення (приросту), призначена для вимірювання складності ЧР. Обчислення ентропії приросту ЕП не потребує жодних припущень про ЧР і може бути застосовним до довільних даних реального світу. Метою [3] є розробка нової систематичної і всеосяжної кількісної основи для оцінки стійкості виробничих процесів, в тому числі представлених ЧР, яка охоплює три аспекти стійкості. Подано рекомендації щодо вибору та кількісної оцінки відповідних показників, перетворення кількісних зважених показників у безрозмірні величини та ранжирування альтернатив розвитку процесів на основі агрегованих балів.

Часова складність відноситься до властивостей ЧР, які є емерджентними, нестійкими або не можуть бути легко описані лінійними процесами [4]. У зазначеній роботі представлений метод *tangle* (клубок), який дозволяє проводити кількісну оцінку складності у відносно коротких ЧР, які зазвичай отримують у багатьох галузях досліджень. Клубок – це міра того, наскільки цей процес відрізняється від простого періодичного руху. Кількісна оцінка методом «Клубок» отримується внаслідок ітеративної процедури масштабування та процедур «розплутування».

В попередніх статтях приблизна ентропія (ПЕ) та ентропія вибірки (ЕВ) використовувалися для кількісної оцінки регулярності ЧР. В [5] було виконано вивчення впливу зміни параметрів моделей ПЕ і ЕВ в часових рядах процесів СОР, а також досліджено здатність цих форм ентропії розрізняти певні групи у ЧР. В [6] гаусову графічну модель (GGM, ненаправлена мережа часткових коефіцієнтів кореляції) докладно описується та досліджується. Встановлюється її значна корисність як інструмента дослідницького аналізу даних. Модель GGM показує, які змінні передбачають один одного, дозволяє розріджене моделювання коваріаційних структур, а також може виявити потенційні причинно-наслідкові зв'язки між змінними, які спостерігаються. Зазначені та значна кількість інших напрямків досліджень процесів, представлених ЧР, свідчать про постійний та значний інтерес до моделей ЧР.

Дослідженням комплексу питань щодо аналізу, класифікації та прогнозування ЧР, в тому числі АЧР процесів залізничного транспорту України, присвячено багато публікацій [11-14]. Серед них виділимо роботи, спрямовані на аналіз первинної природи моделей процесів з ЧР. На рис. 1 зображено загальну структуру процесу відправлення певної категорії вагонів зі станції. На рис. 2, наведені деякі результати моделювання недетермінованих процесів прибуття вагонів на залізничну станцію за рік.

Очевидними для раціонального управління такими процесами є завдання щодо оцінки величин вагонопотоків, які будуть надходити у певний період (наприклад, доба, тиждень ін.). У більшості практичних випадків на залізницях для цього використовують методики нормування. Їх ефективне застосування не завжди є обґрунтованим та ефективним. При цьому важливе значення мають завдання моделювання, аналізу та прогнозування таких процесів на основі ЧР. Для забезпечення достовірності та адекватності математичних моделей і подібних процесів в представлений роботі перш за все було необхідно встановити загальні властивості часових рядів вагонопотоків (ЧР-В).

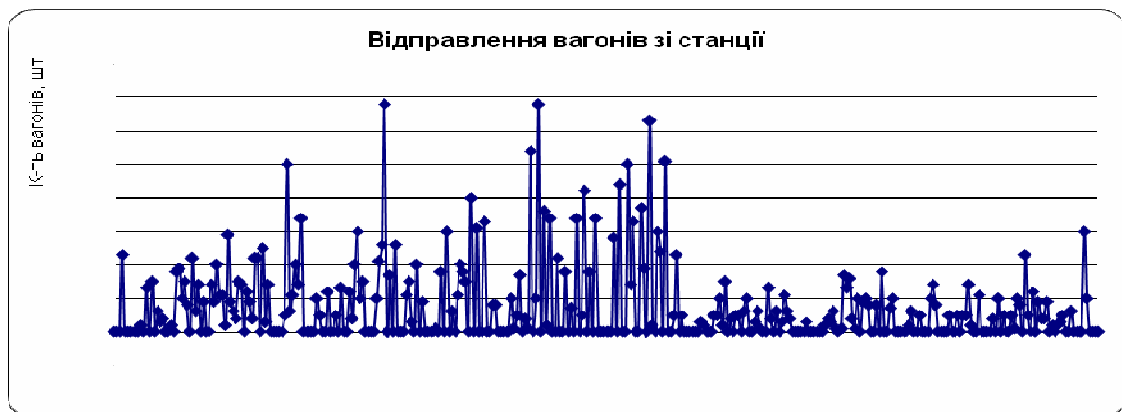


Рисунок 1 – Динаміка відправлення вагонів зі станції за рік



Рисунок 2 – Моделювання процесів прибуття вагонів на станцію

Для цього використовуються методи фрактального аналізу, зокрема, обчислення показника Херста (1) [7, 8]. Його значення дозволяють встановити категорію, класифікувати процеси ЧР на персистентні (трендостійкі), антиперсистентні («злам» тенденції, повернення), виявити випадковий характер процесу.

$$H = \frac{\log(R / S)}{\log(a * N)} \quad (1)$$

де: H – показник Херста; S – середнє квадратичне відхилення ряду спостережень; R – розмах накопиченого відхилення; N – число періодів спостережень; « a » – задана константа.

Відповідно [7] для «коротких» ЧР (відповідають процесам реалізації залізничних перевезень) краще використовувати величину константи $a = \frac{\pi}{2}$. Таке значення параметра (1) дозволяє з більшою вірогідністю оцінювати властивості ЧР і визначати умови, коли ЧР має пам'ять, стає персистентним ($H > 0,5$). Проведені дослідження властивостей часових рядів (ЧР) деяких процесів перевезення та обслуговування вагонопотоків ЗТ України дозволили класифікувати їх як АЧР.

В [10] запропонований теоретико-методичний підхід та спеціалізована процедура, спрямовані на вирішення задачі диференціювання АЧР та їх класифікації. Класифікація АЧР дозволяє виявити додаткові ознаки процесів, за рахунок чого підвищити достовірність і точність результатів аналізу та прогнозування недетермінованих процесів. В основі цього підходу – послідовність перетворень вихідних АЧР за допомогою процедур узагальнення і вирівнювання розташованих поруч рівнів вихідних послідовностей даних. В дослідженнях [7 – 9] були визначені можливості та застосовані методи хаотичної динаміки до аналізу і прогнозування недетермінованих процесів діяльності підприємств ЗТ за умов збурень зовнішнього середовища. При цьому була розроблена методологія та теоретико-методичні підходи до формування завдань класифікації недетермінованих виробничих процесів, представлених АЧР, а також сформовані базові процедури класифікації досліджуваних процесів. З використанням цього інструментарію була отримана класифікація технолого-економічних процесів ЗТ України, представлених ЧР, та запропоновано рекомендації щодо практичного застосування результатів класифікації для підвищення достовірності та точності прогнозування показників процесів.

В статті [8] пропонується наступна процедура класифікації ряду (ПКР) шляхом його перетворення. Схематично за ПКР на основі даних вихідного ЧР

формується серія $ЧР_i(k)$ нових, $k = 2, 3, 4, 5 \dots$. У серії $ЧР_i(k)$ параметр « k » вказує кількість послідовно розташованих рівнів ряду, які використовуються для побудови одного чергового рівня перетвореного ряду (як середнього значення k рівнів) на етапі аналізу « i ». Тобто « i » означає номер етапу ПКР по утворенню та дослідженню $ЧР_i(k)$. На наступному ($i+1$) і подальших етапах процедура ПКР застосовується до нових, утворених на попередніх етапах $ЧР_i(k)$. Побудова серій $ЧР_i(k)$ припиняється, якщо для деякого значення « k » відповідний $ЧР_i(k)$ стане персистентним відповідно (1). В якості іншої, але подібної умови закінчення ПКР, служить виконання вимоги на етапі « i »: $H(ЧР_i(k)) > H^*(1)$. При виконанні умови зупинки серії $ЧР_i(k)$ для декількох узагальнених рівнів « k », вважають, що ЧР відноситься до класу з меншим « k ». До отриманих на етапі « k » модифікованим рядах далі застосовуємо методи кореляційно-регресійного аналізу [9, 10], будують регресійну модель утвореного $ЧР_i(k)$. В якості міри достовірності кожної з моделей виступає коефіцієнт детермінації (R^2). Наведемо приклади застосування процедури ПКР до деяких процесів формування вантажних вагонопотоків ЧР-В. На рис. 3 та рис. 4 представлені узагальнені моделі процесів та регресії для таких модифікованих ЧР-В. Так аналіз дозволив встановити, що вихідний АЧР-В, з моделлю $ЧР_i(5)$, відноситься до 5-го класу, тобто перетворився у персистентний ($H=0,587$) при агрегуванні по 5 послідовних рівнів АЧР. На рис. 5 відображено АЧР-В 7-го класу, тобто він складається з послідовностей модифікованих елементів утворених шляхом агрегації 7-ми ($k=7$) рівнів АЧР.

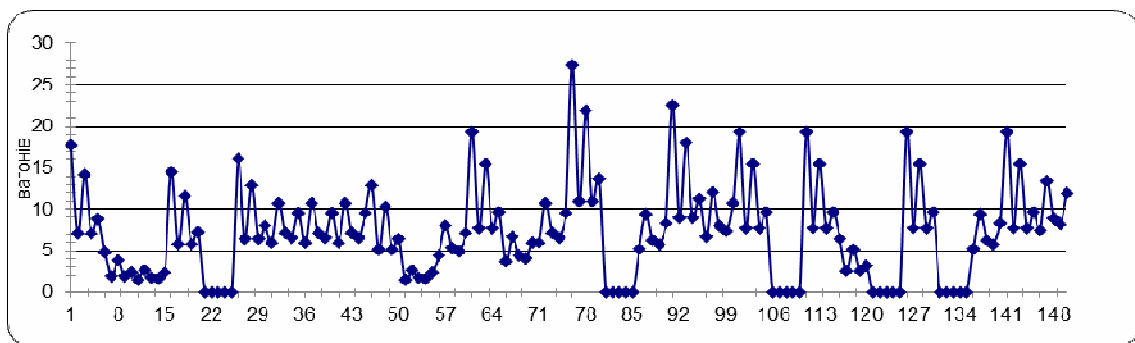


Рисунок 3 - Вихідний АЧР 5-го класу, $ЧР_i(5)$. $H=0,587$

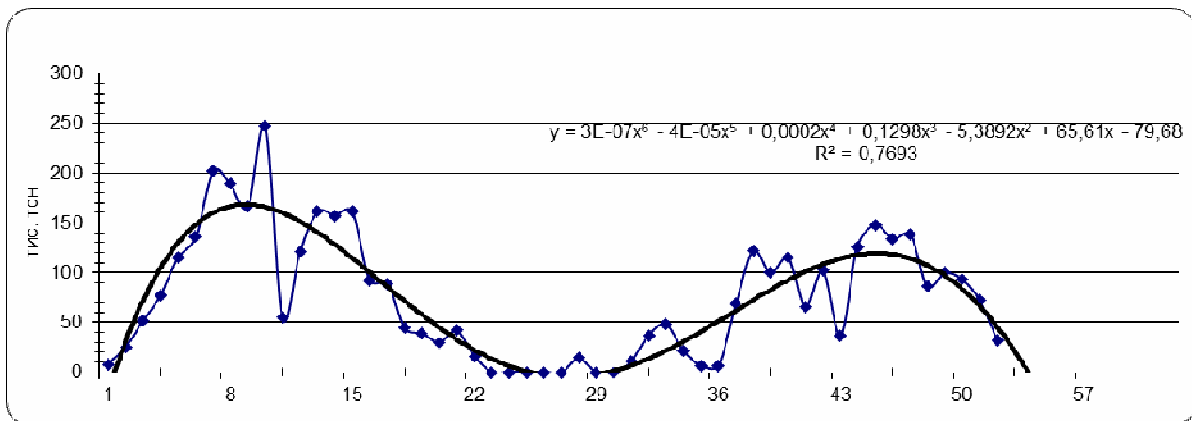


Рисунок. 4 - Модель регресії для $ЧР_i(7)$. ($k=7$)

При цьому поліноміальна модель регресії має досить високий коефіцієнт детермінації ($R^2=0,7693$). Суттєво, що для моделі рис. 4 значення ($k=7$) мало змістовну інтерпретацію, обсяг вагонів за тиждень.

Сформовані на основі модифікованих трендостійких моделей процесів $ЧР_i(k)$, що отримані при класифікації АЧР, забезпечують лише можливість оцінки узагальнених показників за період « k ». Такі значення в багатьох практичних випадках являються недостатніми. В загальному випадку отриманні на основі агрегованого ряду дані не можуть бути використанні безпосередньо, наприклад для оперативного планування. Навіть коли АЧР класифікований як $ЧР_i(k=2)$, на практиці використання такого прогнозу для побудови оперативних планів викликає певні заперечення. Тому бажано мати оцінки рівнів для всіх періодів, а не лише для групових даних. Використання таких прогнозів для ЧР класу $k>2$ ще більш ускладнюється.

Для вирішення завдання оцінювання параметрів процесів на всіх рівнях необхідно застосовувати процедури інтерполяції. Загальна пропозиція щодо форми процедури оцінювання показників рівнів процесів всередині інтервалів класифікації ЧР приведена у [8]. Сутність процедури, призначеної для інтерполяції величин в діапазонах агрегування при класифікації ЧР, полягає в формуванні кількох новоутворених послідовностей класу « k », які зміщені між собою на кілька рівнів. То ж вони мають однаковий клас « k », але відмінну нумерацію рівнів і математичні моделі трендів. За рахунок зміщення моделей можливо виконувати інтерполяцію за рівнями, використовуючи систему рівнянь.

Наведемо приклад розрахунків для $ЧР_i(2)$. Позначим через Δ_1, Δ_2 і т.д. складові першого процесу агрегування (Y_1), $ЧР_1(k_2)$ (Y_1). Рівні моделі процесу (Y_2) $ЧР_2(k_2)(Y_2)$ будуть зміщені від (Y_1) на один рівень. Далі представлені початкові

елементи моделей (Y_1) та (Y_2), елементи яких Δ_1, Δ_2 і т.д., частково містять однакові рівні початкового ЧР.

$$\begin{aligned} Y_1(1) &= \Delta_1 + \Delta_2, \\ Y_2(1) &= \Delta_2 + \Delta_3 \end{aligned} \quad (2)$$

При відомих значеннях першого рівня початкового АЧР та Δ_1 розраховують величину другого рівня процесу (Y_1). Ця величина входить до значення Δ_2 процесу (Y_2), див. рівняння (2). На наступному етапі розрахований параметр за величиною Δ_2 процесу (Y_2), використовують для оцінки рівнів на основі процесу (Y_1). Таким чином поетапно визначають кожен наступний рівень на основі попередніх агрегованих значень моделей (Y_1), (Y_2).

На рис. 5 – 6, наведено приклад моделювання АЧР класу $k=2$ при зазначеному вище початковому зміщенні рівнів. Використання моделей (Y_1) та (Y_2) дозволяє виконувати процедури інтерполяції дозволяє поетапно виконувати процедуру інтерполяції величин всіх рівнів АЧР.

Дослідження отриманих моделей АЧР показали, що модель побудована на основі агрегованих ЧР мають суттєві переваги стосовно точності (абсолютне та відносне відхилення та достовірності (R^2)).

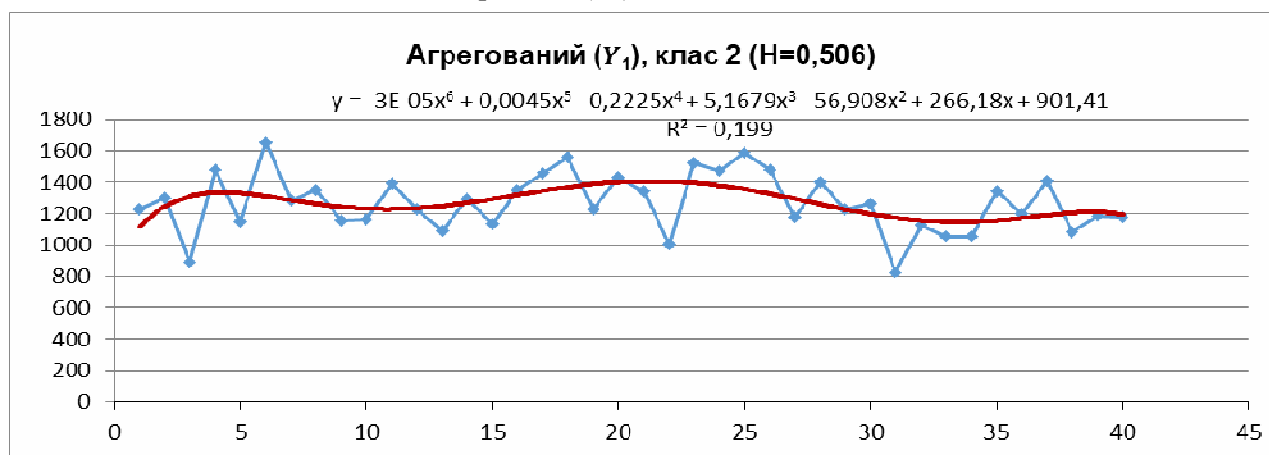


Рисунок 5 - АЧР другого класу $ЧР_i(k=2)$, модель агрегованого ряду Y_1

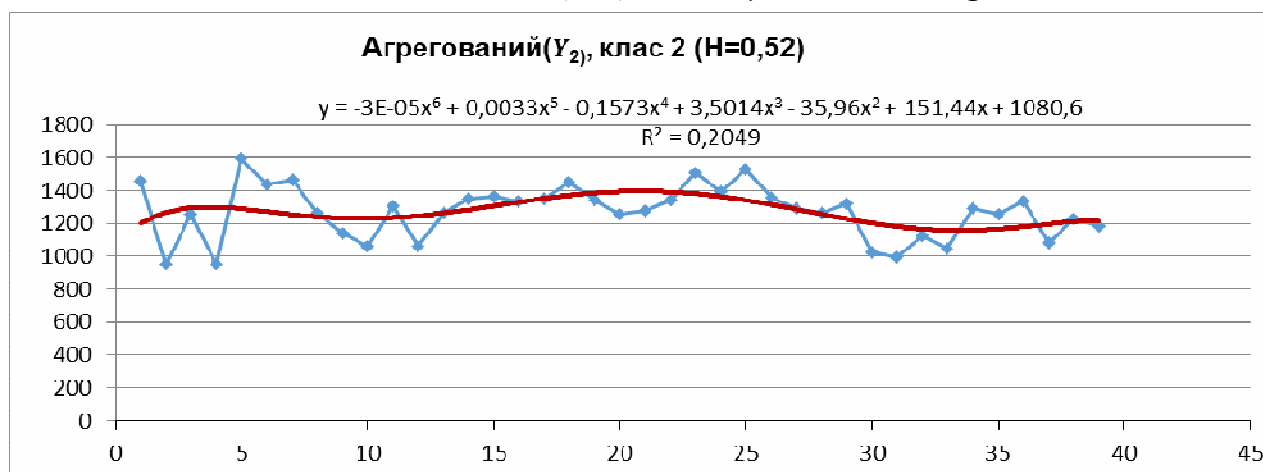


Рисунок 6 - АЧР другого класу $ЧР_i(k=2)$, модель агрегованого ряду Y_2

Для того щоб застосовувати зазначену процедуру інтерполяції до оцінювання та уточнення внутрішніх рівнів ЧР з визначеним класом на основі (2) для ЧР з більш високими класами $k > 2$, необхідно використати наведену нижче систему залежностей, утворених на основі початкового ЧР:

$$\begin{cases} Y_1(1) = \left(\frac{1}{k}\right) \sum_{i=1}^k \Delta_i; \\ Y_2(1) = \left(\frac{1}{k}\right) \sum_{i=2}^{k+1} \Delta_i; \\ \dots \dots \dots \\ Y_k(1) = \left(\frac{1}{k}\right) \sum_{i=k}^{k+k} \Delta_i. \end{cases} \quad (3)$$

Відповідно до (3) для розрахунку оцінок рівнів показника всередині інтервалу агрегування для недетермінованих процесів, представлених числовими послідовностями ЧР класу $k > 1$, необхідно утворити « k » нових послідовностей, елементи яких розраховуються шляхом агрегування та визначення середніх значень послідовностей рівнів у рамках діапазонів індексів. Наступні елементи послідовностей агрегованих моделей розраховують подібно (3), але для інших послідовностей індексів елементів ЧР. Наприклад для $Y_1(2)$ $i=(k+1)$ до $i=2k$. Процедура інтерполяції на основі « k » трендів для зміщених $ЧР_i(k)$ дозволяє отримати прогнозні значення для рівнів АЧР всередині інтервалів агрегування.

Приведені результати досліджень засвідчили значні можливості моделей та процедур агрегування для класифікації та планування антиперсистентних процесів на основі показника Херста. Разом з тим залишились не вирішеними питання щодо вибору обчислювально ефективних схем агрегування рівні АЧР,

а також процедур інтерполяції та короткострокового прогнозування таких недетермінованих процесів.

В статті вирішуються питання, що забезпечують розробку алгоритмічних та інструментальних програмних засобів, а також виконання порівняльного аналізу чисельної ефективності алгоритмів коректного дослідження і класифікації АЧР, формування процедур інтерполяції та короткострокового прогнозування АЧР.

Метою дослідження являється підвищення ефективності та точності методів та алгоритмів класифікації, моделювання та прогнозування антиперсистентних часових послідовностей даних спостережень. Об'єктом дослідження є процеси моделювання та аналізу параметрів недетермінованих часових послідовностей АЧР з рівномірним кроком. Предмет дослідження – процедури чисельного аналізу, класифікації та прогнозування моделей АЧР. Одним із завдань роботи є розробка програмного середовища, для дослідження алгоритмів моделювання антиперсистентних ЧР, оскільки виконаний огляд програмних аналогів показав, що натеper відсутні програмні засоби, які призначені для класифікації та моделювання антиперсистентних часових рядів.

Результати та основний матеріал дослідження. Для обґрунтованого застосування методів фрактального аналізу процесів ЧР, подібних до властивостей ЧР-В рис. 1, рис. 2 ін., у розділі представлені спеціалізовані процедури аналізу, що дозволяють: – 1) однозначно та обчислювально ефективно класифікувати АЧР, диференціювати їх шляхом використання певних схем узагальнення послідовних рівнів початкового ЧР, (класи розрізняються за числом об'єднаних поряд розміщених рівнів, то ж і довжинами інтервалів достовірного планування), – 2) виконувати інтерполяцію значень рівнів показників АЧР в середині інтервалів агрегування, – 3) виконувати процедури короткострокового прогнозування величин сумарних об'ємних показників модифікованого ряду за встановлений інтервал (період узагальнення даних), а також оцінок величин за всіма кроками первинного процесу АЧР.

Дослідження чисельної ефективності та точності алгоритмів класифікації та інтерполяції даних антиперсистентних процесів виконані за методикою порівняльного аналізу. При цьому були сформовані та шляхом чисельного експерименту досліджені варіанти схем узагальнення рівнів, за якими можливо утворювати із початкового ЧР нові агреговані моделі різних класів $ЧР_i(k)$. Так для визначення процесів другого класу $ЧР_i(2)$ із початкової послідовності ЧР:

$$fa : ua_1; ua_2; ua_3 ; ua_4; ua_5; ua_6 ; ua_7; ua_8 ; \quad (5)$$

можливі дві очевидні схеми формування агрегованих моделей:

$$fb(2): (ua_1+ua_2)/2; (ua_2+ua_3)/2; \quad fc(2): (ua_1+ua_2)/2; (ua_3+ua_4)/2; \quad (6)$$

Для процесів третього класу $ЧР_i(3)$ можливі такі схеми формування моделей (з fa):

$$fb(3): (ua_1+ua_2+ua_3)/3; (ua_2+ua_3+ua_4)/3; (ua_3+ua_4+ua_5)/3; \quad (7)$$

$$fc(3): (ua_1+ua_2+ua_3)/3; (ua_3+ua_4+ua_5)/3; (ua_5+ua_6+ua_7)/3; \quad (8)$$

$$fd(3): (ua_1+ua_2+ua_3)/3; (ua_4+ua_5+ua_6)/3; (ua_7+ua_8+ua_9)/3; \quad (9)$$

Для процесів 5-го класу $ЧР_i(5)$ можливі наступні схеми формування (з fa (5)):

$$fb(5): (ua_1+ua_2+ua_3+ua_4+ua_5)/5; (ua_2+ua_3+ua_4+ua_5+ua_6)/5; \quad (10)$$

$$fc(5): (ua_1+ua_2+ua_3+ua_4+ua_5)/5; (ua_3+ua_4+ua_5+ua_6+ua_7)/5; \quad (11)$$

$$fd(5): (ua_1+ua_2+ua_3+ua_4+ua_5)/5; (ua_4+ua_5+ua_6+ua_7+ua_8)/5 \quad (12)$$

та остання:

$$ff(5): (ua_1+ua_2+ua_3+ua_4+ua_5)/5; (ua_6+ua_7+ua_8+ua_9+ua_{10})/5 \quad (13)$$

Зрозуміло що при збільшенні номеру класу АЧР відбувається зростання кількості можливих варіантів схем утворення агрегованих моделей із початкового ЧР. Відзначимо два суттєво відмінні варіанти наведених схем – «без перекриття» членів АЧР ($fc(2)$: (6), $fd(3)$: (9) ін.), коли послідовні елементи агрегованих моделей не містять попередні елементи, а також з перекриттям (повторами) членів ЧР fa в новоутворених агрегованих послідовностях, яким відповідають всі інші наведені моделі $ЧР_i(k)$. В моделях агрегованих процесів (6) – (12) наведено приклади кількох початкових рівнів часових рядів $ЧР_i(k)$.

Приклади схем варіантів агрегування даних (6) – (12), призначених для побудови узагальнених коректних математичних моделей недетермінованих процесів, демонструють змістовність і широкі можливості утворення агрегованих моделей процесів у вигляді $ЧР_i(k)$. При цьому необхідно встановити, по-перше, при якій формі агрегування забезпечується можливість та точність класифікації вихідних ЧР, а по-друге, запропонувати алгоритми прогнозування за агрегованими моделями. Такі алгоритми мають також дозволяти виконувати інтерполяцію – розраховувати внутрішні рівні показників процесів $ЧР_i(k)$, використовуючи оцінки інтегрованих величин, отриманих за моделями відповідних трендів.

Приклади моделей $ЧР_i(k=3)$ при застосуванні процедури агрегування (з перекриттям рівнів), отриманих при дослідженні ЧР (рис. 7). На рис. 7 пред-

ставлений приклад початкового антиперсистентного ЧР (АЧР), що було встановлено за класифікацією Херста (1), з показником Херста $H = 0,479$.

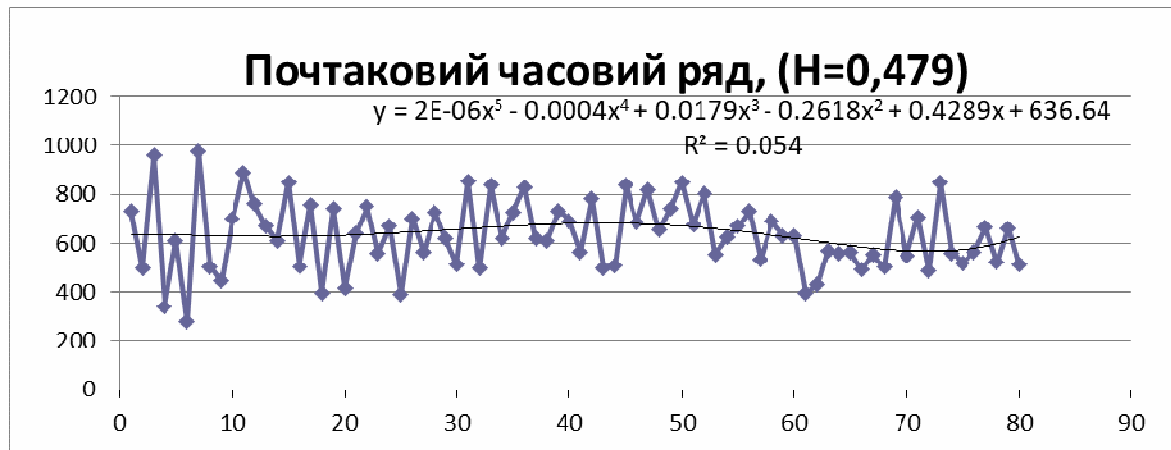


Рисунок 7 – Початковий часовий ряд недетермінованого процесу

За допомогою запропонованої нами процедури агрегування ПКР за раніше наведеною схемою було визначено клас АЧР ($k=3$). При цьому також проведено дослідження різних варіантів схем (7) – (9) формування нових послідовностей ЧР ($k=3$) (з перекриттям та без перекриття рівнів). При цьому утворені наступні агреговані моделі ЧР, фрагменти яких представлені рис. 8 – рис. 11. На рисунках наведені початкові члени нових послідовностей, які відрізняються числом загальних елементів у сформованих рівнях-агрегатах (3).

Зокрема, одним спільним елементом для моделі МА(3_1), двома спільними елементами в моделі МА(3_2), без перетину елементів для МА(3_3). На рисунках також представлено поліноміальні моделі трендів, які розраховані на основі методу найменших квадратів, а також оцінки достовірності поліноміальних моделей (R^2). То ж ці аналітичні моделі представляють узагальнені процеси, що утворені з початкового АЧР.

На основі класифікації за Херстом отримані послідовності (рис.8 – 10) є персистентними, то ж їх застосування при аналізі та прогнозування є припустимим. Використання отриманих математичних моделей агрегованих трендостійких ЧР дає змогу коректно вирішити також завдання інтерполяції, що дозволяє розраховувати оцінки рівнів всередині інтервалів агрегування. Необхідність використання процедур інтерполяції викликана тим, що отримані персистентні моделі агрегованих ЧР дозволяють визначати лише рівні своїх агрегатів. Наприклад, для 3-го класу АЧР ($k=3$) трендостійка модель дасть при аналізі та розрахунку прогнозу сумарне значення для 3-х послідовних рівнів. З метою

практичного використання таких моделей в багатьох випадках необхідно також мати оцінки величин процесу для всіх часових значень.

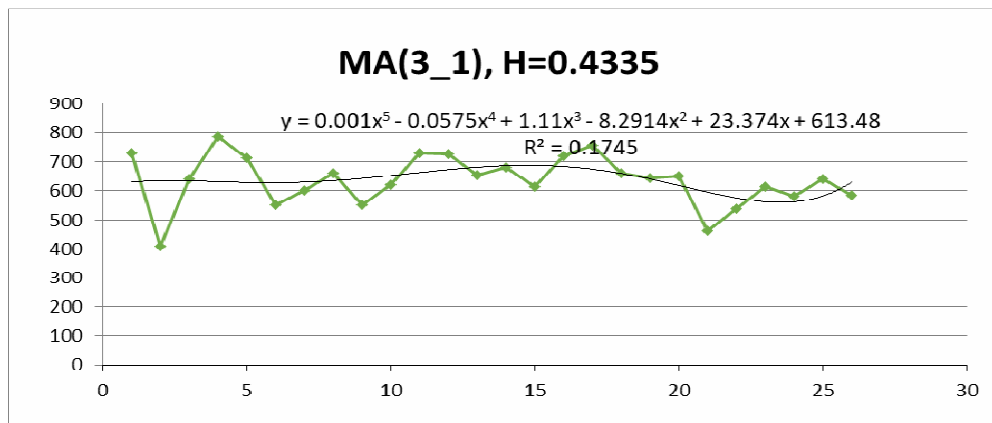


Рисунок 8 – Агрегована модель $ЧР_i(k=3)$ з одним спільним елементом (7)

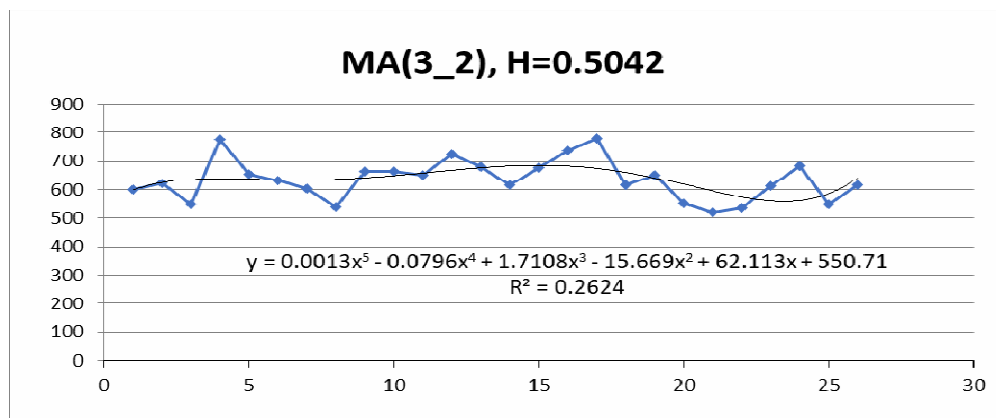


Рисунок 9 – Агрегована модель $ЧР_i(k=3)$ з двома спільними елементами (8)

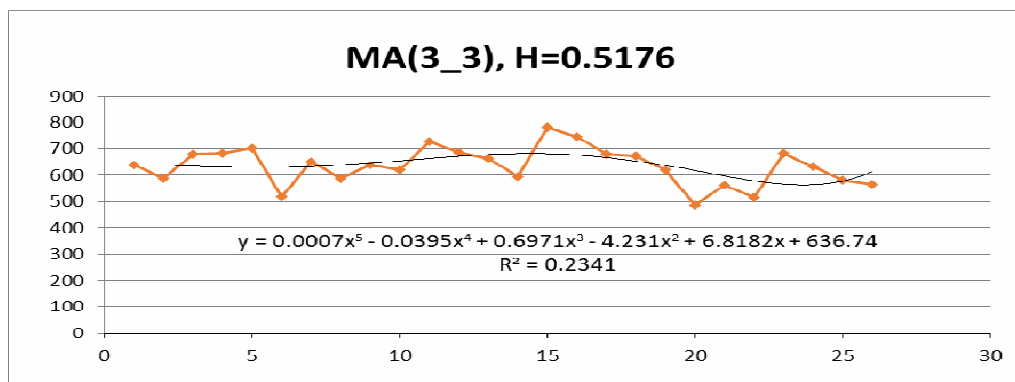


Рисунок 10 – Модель $ЧР$ класу $ЧР_i(k=3)$ для схеми без перекриття рівнів (9), поліноміальна модель та оцінки достовірності (R^2)

У переважній більшості обчислювальних експериментів схеми агрегування з перекриттям рівнів виявилися «нечутливими» до даних при класифікації

ЧР. Майже всі ЧР за такими схемами агрегування мали клас $k=2$. На відміну від цього схеми класифікації на основі агрегування без «перетину» послідовних рівнів дозволили виявити різні класи АЧР. Моделювання та виконаний аналіз показали, що перевагу щодо достовірності класифікації АЧР та точності процедур інтерполяції мають схеми агрегування без перекриття рівнів у послідовних елементах нових утворених ЧР.

Зазначимо можливість використовувати при класифікації АЧР не тільки значення показника Херста $H^*=0,5$ (тобто визначено АЧР(0,5, $k=2$), а й певну іншу, наприклад $H^*=0,6$ (тобто АЧР(0,6, $k=3$)), щоб з більшою «певністю» фіксувати умову закінчення процесу класифікації. Таким чином можуть бути побудовані декілька формальних класифікацій АЧР в залежності від значень граничних величин показника Херста H^* завершення процедур класифікації, або жодної. Остання можливість викликана необґрунтованим вибором значення H^* , що не відповідає процесу ЧР.

Процедури інтерполяції та прогнозування АЧР повинні дозволяти розраховувати внутрішні рівні показників процесів $ЧР_i(k)$, використовуючи оцінки інтегрованих величин, отриманих за моделями відповідних трендів. При цьому для різних варіантів агрегування необхідно було формувати окрему процедуру інтерполяції – розрахунку величин внутрішніх рівнів на основі прогнозних величин для всього діапазону, отриманих на основі системи рівнянь (3). Наприклад, для ($k=3$) за отриманими моделями f_c для розрахування внутрішньо інтервальних значень прогнозних рівнів можливо замість рішення системи рівнянь виду (3) використовувати ($k-1$) початкові значення початкового АЧР. Тоді для розрахунку інтерполяції рівнів достатньо використовувати співвідношення типу (4) замість рішення систем рівнянь:

$$\begin{aligned} Y_1 &= \frac{y_1 + y_2 + y_3}{3}, \text{ то } \rightarrow y_3 = 3Y_1 - y_1 - y_2 \\ Y_2 &= \frac{y_2 + y_3 + y_4}{3}, \text{ то } \rightarrow y_4 = 3Y_2 - y_2 - y_3 \\ Y_3 &= \frac{y_3 + y_4 + y_5}{3}, \text{ то } \rightarrow y_5 = 3Y_3 - y_3 - y_4 \end{aligned} \quad (4)$$

При побудові прогнозу використовується наступний загальний алгоритм: Для процесу Y_1 вибирається k -рівнів початкового ЧР, починаючи з $i=1, n$ по k . Для Y_2 – $i=2, n$ по k , і т.п., для всіх Y_j , $j=1, k$, в залежності від схеми агрегування виду (7) – (12). Визначається коефіцієнт Херста (1), за поліноміальною моделлю $ЧР_i(k)$. будується прогноз та розраховується коефіцієнт детермінації (R^2). За

отриманими моделями розраховуються внутрішньо інтервальні значення прогнозних рівнів, для визначення яких використовуються $(k-1)$ початкові значення ряду Y . Наприклад, для $(k=3)$ виду (13). У дослідженнях функціональної ефективності алгоритмів моделювання $ЧР_i(k)$ показано, що різні схеми агрегування дають різні (хоча і досить схожі) математичні моделі одного й того ж самого процесу, які можна застосувати для оперативного прогнозування в умовах невизначеності. Таким же чином послідовно розраховуються всі наступні рівні прогнозного ряду (рис.11).

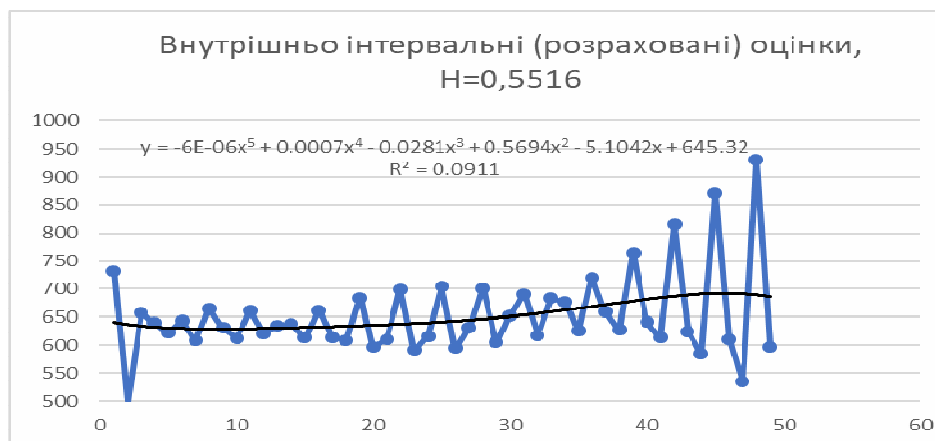


Рисунок 12 – Моделювання прогнозу інтерполяційних оцінок відповідно (4)

Виконані обчислювальні експерименти показали можливість використання запропонованих класифікації $ЧР_i(k)$ лише для короткострокових прогнозів. Для реалізації зазначених алгоритмів інтерполяції та прогнозування рівнів процесів АЧР необхідно мати відповідну структуру їх програмної реалізації, яка забезпечує утворення модельної послідовності що містить всі попередні рівні ЧР.

Інструментальні програмні засоби, створені для дослідження та моделювання наведених вище процесів АЧР відповідно до наведених вище завдань. Діаграма прецедентів (рис. 12) визначає варіанти використання системи аналізу процесів визначених ЧР. На діаграмі представлені наступні функції програми: Ф1 – процедури отримання вхідних даних для дослідження, Ф2 – функція щодо збереження даних та результатів досліджень; Ф3 – відображення графіків моделей процесів; Ф4 – розрахунки оцінок показників Херста для ЧР; Ф5 – побудова класів для моделей за різними схемами агрегування рівнів ЧР; Ф6 – порівняльний аналіз варіантів схем агрегування ЧР; Ф7 – моделювання процесів, представлених ЧР; Ф8 – аналіз моделей процесів ЧР; Ф9 – процедури із формування трендів моделей ЧР; Ф10 – інтерполяція внутрішньо інтервальних даних

для оцінок прогнозів за моделями ЧР; Ф11 – прогнозування процесів за моделями ЧР; Ф12 – процедури оцінювання точності та достовірності (розрахунок R^2) моделей ЧПД; Ф13 – процедури відображення результатів аналізу моделей процесів ЧР; Ф14 – процедури порівняння, виконання аналізу схем агрегування ЧР. В якості базової архітектури програмного комплексу використана архітектура бібліотеки VCL (Visual Component Library).

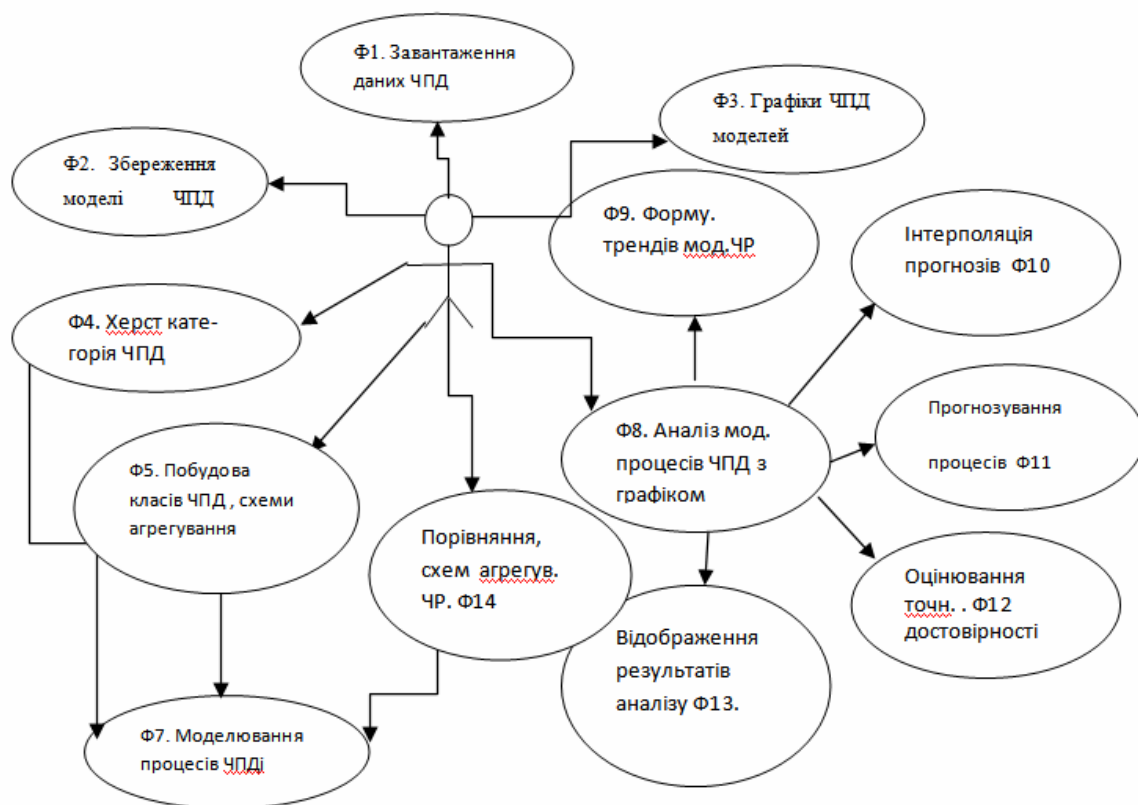


Рисунок 12 – Діаграма варіантів використання при дослідженні ЧР

Висновки. У статті досліджені актуальні завдання щодо класифікації, моделювання і прогнозування АЧР. Наукова новизна результатів роботи визначається наступним. В роботі виконано порівняльний аналіз чисельної ефективності алгоритмів класифікації та короткострокового прогнозування, удосконалено процедуру інтерполяції рівнів агрегованих трендостійких моделей, що забезпечує коректність результатів комп'ютерного моделювання та підвищення точності короткострокового прогнозування показників недетермінованих процесів.

Встановлено, що найбільше детальною та сталою являється класифікація АЧР на основі схем агрегування даних без перетину рівнів. Практичне значення мають результати проведеного порівняльного аналізу чисельної ефективності алгоритмів класифікації АЧР, рекомендації стосовно процедур їх моделю-

вання, а також інструментальні програмні засоби, призначені для автоматизації аналізу АЧР.

ЛІТЕРАТУРА

1. Zhou, Renjie & Yang, Chen & Wan, Jian & Zhang, Wei & Guan, Bo & Xiong, Naixue. (2017). Measuring Complexity and Predictability of Time Series with Flexible Multiscale Entropy for Sensor Networks. *Sensors*. 17. 787. 10.3390/s17040787.
2. Liu, Xiaofeng & Jiang, Aimin & Xu, Ning & Xue, Jianru. (2015). Increment Entropy as a Measure of Complexity for Time Series. *Entropy*. 18. 10.3390/e18010022.
3. Nazzal, Mohammad & Saad, Mohammed. (2018). A general framework for sustainability assessment of manufacturing processes. *Ecological Indicators*. 97. 10.1016/j.ecolind.2018.09.062.
4. Moulder, Robert & Daniel, Katharine & Teachman, Bethany & Boker, Steven. (2019). Tangle: A Metric for Quantifying Complexity and Erratic Behavior in Short Time Series. 10.31234/osf.io/23csa.
5. Montesinos, Luis & Castaldo, Rossana & Pecchia, Leandro. (2018). On the use of approximate entropy and sample entropy with centre of pressure time-series. *Journal of NeuroEngineering and Rehabilitation*. 15. 116. 10.1186/s12984-018-0465-9.
6. Epskamp, Sacha & Waldorp, Lourens & Möttus, René & Borsboom, Denny. (2018). The Gaussian Graphical Model in Cross-Sectional and Time-Series Data. *Multivariate Behavioral Research*. 53. 1-28. 10.1080/00273171.2018.1454823.
7. Скалозуб В.В., Клименко І. В. Розвиток процедур аналізу та прогнозування недетермінованих технологіко-економічних процесів на основі показників хаотичної динаміки. *Економіка: реалії часу. Науковий журнал*. – 2016. – № 4 (26). – С. 149-154. Цитується у базах Google Scholar, Index Copernicus.
8. Скалозуб В. В., Клименко І. В. Метод планирования недетерминированных процессов эксплуатации парка железнодорожных технических систем. *Наука та прогрес транспорту. Вісник Дніпропетровського національного університету залізничного транспорту*, 2018, № 5 (77). – С. 7-18.
9. Скалозуб В. В., Жуковичкий І. В, Клименко І. В, Заец А.П..Создание интеллектуальных систем поддержки принятия решений в единой автоматизированной системе управления грузовыми железнодорожными перевозками Украины. *Системні технології : регіон. міжвуз. зб. наук. пр. – Дніпро*, 2018. – № 3 (116). – С. 153–162.
10. Скалозуб В. В., Клименко І. В. Оценка и прогнозирование параметров временных рядов вагонопотоков на основе специализированной процедуры фрактального анализа. *Наук. вісн. Херсон. держ. ун-ту. Серія : Економічні науки*. – 2016. – Вип. 20, ч. 2. – С. 189-192.
11. Скритник О.Я. Фрактальний аналіз часових рядів даних спостережень за температурою повітря в Україні / О.Я. Скритник, О.А. Скритник, Д.О. Ошурко // *Гідрологія, гідрохімія і гідро екологія*. – 2013. Т. 2. (29). – С. 89 – 95

12. Бойко Р.О., Загоровська Л.Г., Ярова Т.В. Інформаційне забезпечення задачі прогнозування техніко-економічних показників функціонування технологічного комплексу. Східно-європейський журнал передових технологій. – Харків, 2013. – Том 3. № 10 (63). – С. 62–64.
13. Долгіх, А.О., Байбуз, О.Г. Порівняльний аналіз ефективності моделей прогнозування часових рядів за допомогою багатокритеріальної процедури на прикладі фінансових показників. Вісник Житомирського державного технологічного університету. Серія: Технічні науки. – Житомир, 2019. – 83(1), С.131-141.
14. Кораблев, Н.М., Иващенко Г.С. Краткосрочное прогнозирование временных рядов, содержащих аномальные значения, при помощи моделей искусственных иммунных систем. Бионика интеллекта – Харків, 2013. – 2(85), С.95-99.

REFERENCES

1. Zhou, Renjie & Yang, Chen & Wan, Jian & Zhang, Wei & Guan, Bo & Xiong, Naixue. (2017). Measuring Complexity and Predictability of Time Series with Flexible Multiscale Entropy for Sensor Networks. *Sensors*. 17. 787. 10.3390/s17040787.
2. Liu, Xiaofeng & Jiang, Aimin & xu, Ning & Xue, Jianru. (2015). Increment Entropy as a Measure of Complexity for Time Series. *Entropy*. 18. 10.3390/e18010022.
3. Nazzal, Mohammad & Saad, Mohammed. (2018). A general framework for sustainability assessment of manufacturing processes. *Ecological Indicators*. 97. 10.1016/j.ecolind.2018.09.062.
4. Moulder, Robert & Daniel, Katharine & Teachman, Bethany & Boker, Steven. (2019). Tangle: A Metric for Quantifying Complexity and Erratic Behavior in Short Time Series. 10.31234/osf.io/23csa.
5. Montesinos, Luis & Castaldo, Rossana & Pecchia, Leandro. (2018). On the use of approximate entropy and sample entropy with centre of pressure time-series. *Journal of NeuroEngineering and Rehabilitation*. 15. 116. 10.1186/s12984-018-0465-9.
6. Epskamp, Sacha & Waldorp, Lourens & Möttus, René & Borsboom, Denny. (2018). The Gaussian Graphical Model in Cross-Sectional and Time-Series Data. *Multivariate Behavioral Research*. 53. 1-28. 10.1080/00273171.2018.1454823.
7. Skalozub V.V., & Klymenko Y.V. (2016). The development of analysis and forecasting procedures of non-deterministic technological and economic processes on the basis of indicators of chaotic dynamics. *Economics: Time Realities*, 4(26), 82-90.
8. Skalozub, V. V., & Klymenko, Y. V. (2018) Method for planning non-determined operation processes of railway technical system park. *Nauka ta prohres transport. Visnyk Dnipropetrovskoho natsionalnoho universytetu zaliznychnoho transportu imeni akademika V. Lazariana*, 5(77), 7-18.
9. Skalozub, V. V., Zhukovitskiy, I. V., Klimenko, I. V., & Zaets, A. P. (2018). Creation of Intellectual Decision Support Systems in a Unified Automated System for

Managing Rail Freight in Ukraine. Systemni tekhnolohii: Rehionalnyi mizhvuzivskiyi, 3(116), 153-162.

10. Skalozub V.V. & Klymenko I.V. (2016) Estimation and Prediction of the Parameters of Time Series of Traffic Volumes Based on the Specialized Procedures of the Fractal Analysis. Naukovyi visnyk Khersonskoho derzhavnoho universytetu, 20 (II), 189-192.

11. Skrynyk, Oleg & Skynyk, Olesya & Oshurok, Dmytro. (2013). Fractal analysis of air temperature time series collected in Ukraine. Hidrolohiia, hidrokhimiia i hidroekolohiia. 2(29). 89-95.

12. Boyko, R. O., Zagorovska, L. G., & Yarova, T. V. (2013) Informative Providing of Task of Prognostication of Feasibility Indicators of Functioning of Technological Complex. Eastern-European Journal of Enterprise Technologies, 3(10), 62-64.

13. Dolgih, A.O. and Bajbuz, O.G. (2019) Comparative analysis of time series forecasting models effectiveness by means of a multi-criteria procedure on the example of financial indicators. "The Journal of Zhytomyr State Technological University"? Section 1: Engineering, 1(83), 130-141.

14. Korablev, N.M, and Ivashchenko G.S. (2015) Application of the models of artificial immune systems for short-term forecasting of time series containing abnormal values. Bionics of Intelligense: Sci. Mag, 2 (85), 95-99.

Received 30.03.2022.

Accepted 01.04.2022.

***Models and procedures for classification and forecasting
of nondeterministic processes according to chaotic dynamics parameters***

The article investigates the processes of classification, modeling and short-term prediction of nondeterministic time sequences, which are represented by antipersistent time series (ATS). The subject of analysis - procedures for classification and forecasting the parameters of such models. The object of research is the processes of modeling and analysis of parameters of nondeterministic time series of ATS with a uniform step. The aim of the work is to increase the efficiency and accuracy of methods and algorithms for classification, modeling and forecasting of ATS. Models and methods of fractal analysis are used to study the properties of ATS, on the basis of which the categories of processes of numerical series are established. With the help of aggregation of ATS levels correct mathematical models of classification of nondeterministic time sequences are developed, and also algorithmic and software means of their realization are formed. Examples of models of numerical series obtained using the aggregation procedure presented in the study are given. It is established that the most detailed and stable is the classification of ATS based on data aggregation schemes without level crossing. The comparative analysis of numerical efficiency of algorithms of classification of ATS is carried out and the task of formation of procedures of interpolation and short-term forecasting of ATS is realized. An instrumental software environment is presented, which provides a correct study of algorithms for modeling and classification of antipersistent time series. Recommendations on the procedures for modeling ATS classification algorithms are of practical importance.

Скалозуб Владислав Васильович - професор, каф. «Комп'ютерні інформаційні технології», Український державний університет науки і технологій, м. Дніпро.

Горячкін Вадим Миколайович – зав. кафедрою «Комп'ютерні інформаційні технології», Український державний університет науки і технологій, м. Дніпро.

Клименко Іван Вікторович – ас. кафедри «Комп'ютерні інформаційні технології», Український державний університет науки і технологій, м. Дніпро.

Шваповал Данило Олегович – магістр кафедри «Комп'ютерні інформаційні технології», Український державний університет науки і технологій, м. Дніпро.

Skalozub Vladislav – professor, Department “Computer and Information Technology”, Ukrainian State University of Science and Technologies, Dnipro.

Horiachkin Vadim – head of Department “Computer and Information Technology”, Ukrainian State University of Science and Technologies, Dnipro.

Klymenko Ivan – assistants of Department “Computer and Information Technology”, Ukrainian State University of Science and Technologies, Dnipro.

Shapoval Danylo – magister of Department “Computer and Information Technology”, Ukrainian State University of Science and Technologies, Dnipro.