УДК 681.3:664.1

В.І. ЗАЇКА, асп.
В.Д. КИШЕНЬКО, канд. техн. наук
Національний університет харчових технологій

ЗАСТОСУВАННЯ АПАРАТУ НЕЛІНІЙНОГО АНАЛІЗУ ДИНАМІЧНИХ СИСТЕМ ДЛЯ ОБРОБКИ ІСТОРИЧНИХ ДАННИХ РОБОТИ СТАНЦІЇ ДЕФЕКОСАТУРАЦІЇ

У даній статті розглядається аналіз часових рядів за допомогою алгоритмів нелінійної динаміки: відновлення фазового простору, показника Херста, кореляційної розмірності, відновлення атрактора динамічної системи.

Ключові слова: *фазовий простір, показник Херста, кореляційна розмірність.*

В данной статье рассматривается анализ временных рядов с помощью алгоритмов нелинейной динамики: восстановления фазового пространства, показателя Херста, корреляционной размерности, восстановления аттрактора динамической системы.

Ключевые слова: *фазовое пространство, показатель Херста, корреляционная размерность.*

This article discusses the time series analysis using algorithms of nonlinear dynamics: the restoration of the phase space, Herst exponent, correlation dimension, reconstruction of the attractor of a dynamical system.

Keywords: *phase space, Herst index, the correlation dimension.*

Традиційні моделі управління технологічними процесами і традиційніаналітичні методи аналізу показників ефективності і прогнозування все частіше і частіше натикаються на проблеми, що не мають ефективного вирішення в рамках відомих рішень. Традиційні підходи, що стали вже класичними, були розроблені для опису стійких і не радикально змінних процесів — які не сильно відхиляються від стану рівноваги. По самій своїй суті ці методи і підходи не були призначені для опису і моделювання швидких змін, непередбачених стрибків і складних взаємодій окремих складових технологічного процесу.

Зміни, які відбуваються на станції дефекосатурації, мінливі настільки інтенсивно, а їхні якісні показники бувають настільки несподіваними, що для аналізу і прогнозування вихідних параметрів, тобто якості очищення дифузійного соку, синтез нових аналітичних і обчислювальних підходів, що беруть свій початок у різних областях людських знань, став насущною практичною необхідністю. Необхідно вивчати динамічні процеси, що відбуваються в незворотних багатокомпонентних інтерактивних адаптивних системах, розглядати причини та механізми виникнення нових режимів і структур, вивчати характерні масштаби і швидкості перехідних та сталих процесів, передбачувати ймовірні зміни системи з метою забезпечення можливості управління несподіваними кардинальними змінами в динамічному режимі, що виникають у складних системах. У теорії складних систем досліджуються, головним чином, нелінійні системи зі зворотним зв'язком, коли інформація з виходу системи подається на вхід і стає наступним набором вхідних даних. Технологічний процес дефекосатурації можна віднести до таких систем. Останні роки ознаменувалися підвищеним інтересом до пошуку нелінійних моделей, які могли б адекватно відтворювати складні патерни динамічних процесів, оскільки вже стало ясно, що лінійний підхід до аналізу вихідних параметрів не дозволяє змоделювати сильно нерегулярне поводження, характерне для більшості технологічних систем [1]. Існує кілька конкуруючих підходів, що використовують ідею нелінійності. Традиційні моделі є стохастичними. Однак ті обмеження, які використовуються при побудові моделі з метою зробити її придатною для практичного використання, по суті справи, знищують внутрішню «складність», яка властива динамічному процесу [2].

У зв'язку із цим, останнім часом інтенсивно розвивається альтернативний підхід до аналізу нелінійностей, а саме підхід, що базується на теорії детерміністичного хаосу, що пропонує пояснення іррегулярному поводженню і аномаліям у системах, які, не є по своїй природі стохастичними. Теорія хаосу пропонує зовсім нові концепції і алгоритми для аналізу часових рядів, які можуть привести до більше глибокого і повного розуміння відображуваних ними процесів.

У роботі ми зупинимося лише на аспекті застосування теорії складних систем в аналізі такого вихідного показника технологічного процесу, як якісного показника роботи станції дефекосатурації, величини pH соку: методах хаотичної динаміки.

В основу сучасних методів аналізу покладений пошук моделей нелінійного поводження часових рядів. Це пояснюється тим, що нелінійні моделі можуть уловлювати дуже складні патерни в часових рядах.

Часовим рядом зазвичай називається послідовність подій, спостережуваних через деякі, як правило, рівні інтервали часу. Стосовно до технології очищення дифузійного соку цим параметром може бути величина pH, концентрація сухих речовин, витрата очищеного соку.

Головна ідея застосування методів хаотичної динаміки до аналізу часових рядів полягає в тому, що основна структура хаотичної системи, що містить у собі всю інформацію про систему, а саме атрактор динамічної системи (підмножина фазового простору, що притягує траєкторії в межі нескінченного часу), може бути відновлена через вимірювання тільки однієї спостережуваної характеристики цієї динамічної системи, фіксованої як часовий ряд. Відповідно до методу Грасбергера і Прокаччі процедура реконструкції фазового простору та відновлення хаотичного атрактора системи при динамічному аналізі часового ряду зводиться до побудови так званого фазового простору.

Для динамічного аналізу даних використовувалися наступні програмні продукти: програма Dataplore компанії DATAN, набір модульних програм TISEAN, створених в Інституті фізики й теоретичної хімії університету міста Франкфурта Райнером Хеггером (Rainer Hegger), Ольгером Канцом (Holger Kantz) і Томасом Шрайбером (Thomas Schreiber), Chaos Data Analyzer (CDA), написана Дж. Спроттом і Дж. Роуландсом (Sprott J., Rowlands G.), а також програма Fractan, розроблена в інституті математичних проблем біології РАН В. Сичьовим. Для фрактального аналізу в роботі використовувався спеціалізований набір розширень (toolbox) програми MatLab - FracLab, створений у французькій лабораторії Fractals team.

Головна ідея застосування методів хаотичної динаміки до аналізу часових рядів полягає в наступному. Виявляється [3], основна структура хаотичної системи, що містить у собі всю інформацію про систему, а саме, її атрактор (підмножина фазового простору, що притягає траєкторії в межі нескінченного часу), може бути відновлена через вимірювання тільки одного спостережуваного параметру цієї динамічної системи, фіксованого як часовий ряд.

Детерміновані динамічні системи описують еволюцію системи із часом у деякому фазовому просторі

$$\Gamma \subset R^d \quad (1)$$

Ці системи можуть бути породжені, наприклад, звичайними диференціальними рівняннями

$$x = F(x(t)) \quad (2)$$

або, якщо час дискретно $t = n\Delta t$, виразом виду

$$x_{n+1} = f(x_n) \quad (3)$$

Часові ряди потім можна розглядати як послідовність спостережень $\{S_n = S(x_n)\}$. Тому що послідовність (звичайно скалярна) $\{S_n\}$ сама по собі не породжує багатомірний фазовий простір динамічної системи, необхідно використовувати деякий технічний прийом, щоб розкрити багатомірну структуру, використовуючи тільки наявні дані.

Відповідно до методу Грасбергера і Прокаччі [4], процедура реконструкції фазового простору і відновлення хаотичного атрактора системи при динамічному аналізі часового ряду, зводиться до побудови так званого лагового або відновленого простору за допомогою методу затримки (method of delays). Вектори в новому просторі, просторі вкладення, сформовані зі значень часового ряду скалярних вимірів з часовим запізнюванням:

$$\overline{S}_n = (S_{n-(m-1)r}, S_{n-(m-2)r}, \ldots, S_n) \quad (4)$$

Число елементів m називається розмірністю вкладення, час ф зазвичай називається затримкою або лагом. У теоремах Такенса [5] і Сойера [6] показується, що якщо послідовність $\{S_n\}$ насправді складається зі скалярних вимірювань структури динамічної системи, тоді, при певних припущеннях, таке відновлення фазового портрета є точною картиною справжньої множини {x}, якщо, m досить велике. Або, інакше кажучи, реальний атрактор динамічної системи і «атрактор», відновлений у лаговому просторі по часовому ряді відповідно до зазначеного вище правила (псевдоатрактор), при адекватному підборі розмірності вкладення m, є топологічно еквівалентними.

Для обчислення характеристик псевдоатрактора, таких як його фрактальна розмірність, показники Ляпунова, необхідно мати множину точок, визначених у фазовому просторі розмірності m і належних атрактору. Число точок M у розрахунках кінцеве, але повинно бути досить великим. Відповідно до формули, запропонованої в [7]

$$M \geq M_{min} = 10^{2+0.4D} \quad (5)$$

де D — розмірність атрактора. Виходячи з цього, нами була обрана така кількість даних, що цілком відповідала б малорозмірній динаміці руху, з розмірністю n від 4 до 8. Відповідно до цих міркувань було прийнято рішення вибрати довжини часових рядів ≈ 16000 вимірювань.

Традиційний спосіб вибору часової затримки складається в обчисленні автокореляційної функції часового ряду:

$$A(\tau) = \frac{1}{m}\sum_{k=0}^{m-1}(S_k - \overline{S})(S_{k+\tau} - \overline{S});$$
$$m = M - \tau \quad (6)$$

Затримка ф вибирається рівною часу першого перетинання нуля автокореляційної функції. Інший спосіб вимагає обчислення спектра потужності часового ряду, тобто швидкості перетворення Фур'є автокореляційної функції. Якщо в спектрі потужності присутні кратні піки, то затримка ф вибирається рівна чверті періоду найвищої з домінуючих частот. Третій спосіб заснований на обчисленні функції затриманої взаємної інформації (time delayed mutual information). Функція взаємної інформації S визначається за формулою:

$$S = -\sum_{ij} p_{ij}(\tau)\ln\frac{p_{ij}(\tau)}{p_i p_j} \quad (7)$$

де для деякої розбивки даних p_i — імовірність знайти значення часового ряду в i-му інтервалі розбивки, а P_{ij}(ф) — спільна ймовірність того, що спостережувана величина потрапить в i-й інтервал, а час цього спостереження далі потрапить в j-й інтервал.

Ідея програмної реалізації методу наступна. Спочатку відновлюється фазовий простір за допомогою методу Грасбергера і Прокаччі. Вектори в цьому просторі визначаються відповідно до формули (3). Для кожної точки S_i часового ряду, шукається його найближчий сусід S_j в m-мірному просторі. Обчислюється відстань між ними

$$\| \overline{S}_i - \overline{S}_j \| \quad (8)$$

Далі, інтегруємо обидві точки і обчислюємо

$$R_i = \frac{| \overline{S}_{i+1} - \overline{S}_{j+1} |}{\| \overline{S}_i - \overline{S}_j \|} \quad (9)$$

Якщо R_i перевищує даний евристичний поріг R_t, то вважається, що точка має помилкового найближчого сусіда.

Один з тестів, застосовуваних на практиці для перевірки наявності хаотичної складової в досліджуваному ряді фінансових даних, складається у вивченні властивостей кореляційної суми Cm(r) і поводження кореляційної розмірності Dm(r) залежно від розмірності вкладення m. Кореляційна сума Cm(r) — це ймовірність того, що пари точок на відновленому атракторі в m-мірному лаговому просторі перебуває в межах відстані r один від одного. Якщо графік функції ln Cm(r) відносно ln r має чітко виражену лінійну ділянку, це вказує на існування самоподібної геометрії атрактора, що, у свою чергу, говорить про хаотичність процесу.

У випадку стилізованих даних, коли нам відома розмірність n фазового простору динамічної системи і всі n координати кожної точки на атракторі, кореляційну розмірність D2 атрактора знаходять у такий спосіб: розглядається кореляційний інтеграл C(r), що показує відносне число пар точок атрактора, що перебувають на відстані, не більшому r:

$$C(r)=\frac{2}{m(m-1)}\sum_{i=0}^{m-2}\sum_{j=i+1}^{m-1}\theta(r-p(x_i,x_j)) \quad (10)$$

де θ — функція Хевісайда, p — відстань в n — мірному фазовому просторі, m — число точок x_i на атракторі. На досить малих масштабах довжин і коли розмірність вкладення m не менше топологічної розмірності атрактора, виконується залежність:

$$C(r)\to r^{D_2} \quad (11)$$

де D_2 — шукана кореляційна розмірність атрактора. Про логарифмувавши рівняння (11):

$$\ln C(r)\to D_2 \ln r \quad (12)$$

Вираз (12) дає шукану оцінку розмірності атрактора, як тангенса кута нахилу прямій, що апроксимує графік кореляційного інтеграла C(r) у подвійному логарифмічному масштабі.

Обчислення показника Херста проводиться за допомогою методу нормованого розмаху (Rescaled Range (R/S) Analysis). Основною метою обчислення показника Херста є визначення довгострокової кореляції в часовому ряді, і виявлення його фрактальної структури. Крім того, відзначимо, що за допомогою R/S-аналізу можна виявити існуючі в динаміці системи статистичні цикли [8].

У відповідності зі значенням показника Херста H, всі часові ряди можуть бути класифіковані на три типи [9]:

антиперсистентні часові ряди ($0 < H < 0.5$);

випадкові часові ряди ($H = 0.5$);

персистентні часові ряди ($0.5 < H < 1$).

Дослідження значень показника Херста проводилися в програмному середовищі — Fractan.

Результати дослідження R/S — аналізу даних наведені в таблиці 1.

Таблиця 1

Показник Херста для величини pH.

Доба роботи	Показник Херста, H	Фрактальна розмірність, D
1	1.0775	0.9225
2	0.7435	1.2565
3	0.6403	1.3597
4	1.0639	0.9361
5	0.6802	1.3198
6	0.6397	1.3603
7	0.610	1.3990
8	1.0647	0.9353
9	0.8445	1.1555
10	1.1641	0.8359

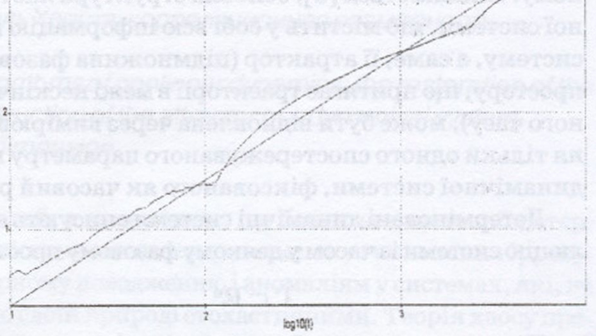

Рис 1. R/S — аналіз досліджуваних даних величини pH

Виходячи з отриманих результатів можна зробити висновок, що характер процесів, які проходять на станції дефекосатурації має в основному персистентний вид ($0.5 < H < 1$).

Вище було відзначено, що визначення кореляційної розмірності часового ряду даних є одним з основних тестів, застосовуваних на практиці для з'ясування наявності в них хаотичної складової. Оцінка кореляційної розмірності повинна проводиться для логарифмів даних [6]. Серед всіх програмних продуктів, використаних у ході досліджень, процедуру оцінки кореляційної розмірності виконують лише дві програми: CDA і Fractan.

Результати дослідження даних наведені в табл. 2.

Таблиця 2

Залежність кореляційної розмірності від розмірності вкладення

Доба роботи ДС	Кореляційна розмірність, D_2	Розмірність фазового простору, n
1	4.563	5
2	5.468	7
3	3.233	6
4	2.334	4
5	2.474	7
6	2.205	4
7	2.902	7
8	0.185	1
9	1.437	3
10	5.712	7

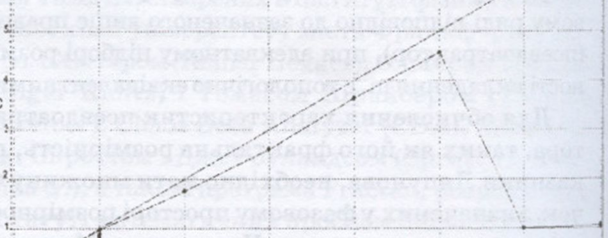

Рис 2. Залежність кореляційної розмірності від розмірності вкладення для ряду величини pH

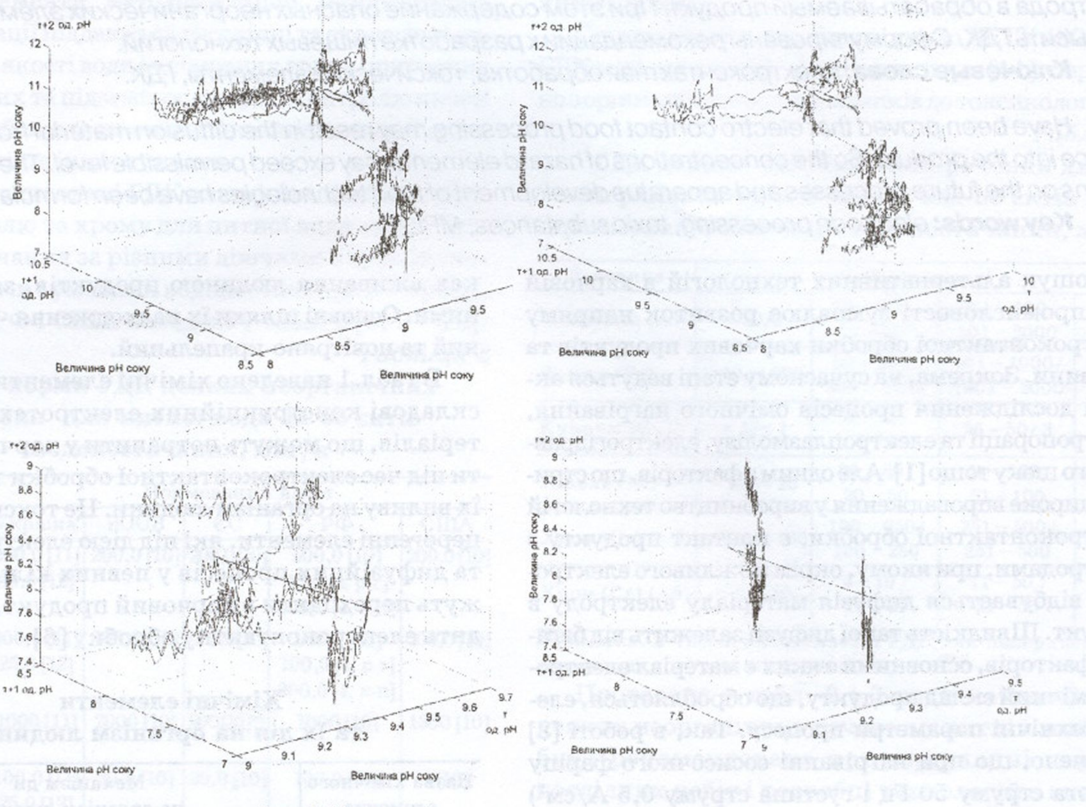

Рис. 3 Відновлені атрактори часових рядів величини pH дифузійного соку II-ї станції дефекосатурації

Відновлені атрактори часових рядів величини pH дифузійного соку роботи станції дефекосатурації наведені на рис.3.

ЛІТЕРАТУРА

1. *Малинецкий Г.*, Потапов А., Современные проблемы нелинейной динамики //М: Эдиториал УРСС, 2000. — 328 с.

2. *Кремер Н.* Теория вероятностей и математическая статистика. // М.: Юнити, 2000. — 573 с.

3. *Куперин Ю.А.*, Дмитриева Л.А., Сорока И.В., Методы теории сложных систем в экономике и финансах.// URL, http://is2001.icape.ru/thesis/7.html

4. *Grassberger P.*, Procaccia I. Characterization of strange attractors. // Phys. Rev. Lett. 50, 346-349 (1983).

5. *F. Takens*, Detecting Strange Attractors in Turbulence // Lecture Notes in Math. Vol. 898, Springer, New York (1981).

6. *T. Sauer*, J. Yorke, M. Casdagli, Embedology // J. Stat. Phys. 65, 579 (1991).

7. *Nerenberg M.A.*, Essex C. Correlation dimension and systematic geometric effects.// Phys.Rev. A 42, 7605 (1986).

8. *Grassberger P.*, Procaccia I. Characterization of strange attractors. // Phys. Rev. Lett. 50, 346-349 (1983).

9. *Петерс Э.* Хаос и порядок на рынках капитала. // М: Мир, 2000. — 305с.

Одержана редколегією 09.03.2010