

3. Карев, В. М. Механика гидрогазоимпульсного воздействия на трещиновато-пористую породу при скважинной гидродобыче [Текст] / В. И. Карев, Ю. Ф. Коваленко, В. Н. Одинцев // Физико-технические проблемы разработки полезных ископаемых. – 1995. – №6. – С. 70-83.
4. Пучков, Л. А. Извлечение метана из угольных пластов [Текст] / Л. А. Пучков. – М: Из-во МГТУ. – 2002. – 383 с.
5. Фізико-хімічна геотехнологія: навч. посібник [Текст] / М. М. Табаченко, О. Б. Владико, О. Є. Хоменко, Д. В. Мальцев. – Д.: Національний гірничий університет, 2012. – 310 с.
6. Boger, C. Production gas from its source. Oilfield Review [Текст] / Boger C., Kieschnick I., Lewis R. E. Waters G. – Autumn 2006, – P. 36-49.
7. Zhao, C. Finite element modeling of methane gas / Migration in coal seams [Текст] / C. Zhao, S. Villiappan // Computers and structures. – 1995. – V.55, №4. – P. 625-629.
8. Rice, J. R. Some basic stress-diffusion solutions for fluid-saturated elastic porous media with compressible constituents [Текст] / J. R. Rice, M. P. Cleary // Reviews of Geophysics and Space Physics. – 1992. – №14. – P. 227-241.
9. Задачі термодифузії та методи їх розв'язку [Текст] / під ред. д.т.н. В. П. Ляшенко. – Кременчук: КНУ ім. М. Остроградського, 2012. – 112 с.
10. Булгаков, В. К. Моделирование горения полимерных материалов [Текст] / В. К. Булгаков, В. И. Кодолов, А. М. Липанов. – М: Химия, 1990. – 240 с.
11. Дьяченко, В. Ф. Основные понятия вычислительной математики [Текст] / В. Ф. Дьяченко. – М: Наука, 1977. – 128 с.

*Використання ідей генетичних алгоритмів до побудови математичних моделей дає можливість не тільки вибрати оптимальну за структурою адекватну модель, але й значно зменшити число обчислень при переборі моделей. Тому новий метод може бути застосований для прогнозування, як фізичних явищ, так і складних технологічних процесів. Ефективність розробленого методу підтверджена на приклад побудови моделі зміни рівня води у ріці Дністер у залежності від погодних умов*

*Ключові слова: синтез математичних моделей, генетичний алгоритм, хромосома, зовнішній критерій, некротні частоти, гармонічний ряд, точність прогнозу*

*Использование идей генетических алгоритмов к построению математических моделей дает возможность не только выбрать оптимальную по структуре адекватную модель, но и значительно уменьшить число вычислений при переборе моделей. Поэтому новый метод может быть применен для прогнозирования, как физических явлений, так и сложных технологических процессов. Эффективность разработанного метода подтверждена на примере модели изменения уровня воды в реке Днестр в зависимости от погодных условий*

*Ключевые слова: синтез математических моделей, генетический алгоритм, хромосома, внешний критерий, некротные частоты, гармонический ряд, точность прогноза*

УДК 628.5.66

## МЕТОД ПРОГНОЗУВАННЯ РІВНЯ ВОДИ У Р. ДНІСТЕР У ЗАЛЕЖНОСТІ ВІД ПОГОДНИХ УМОВ

**М. І. Горбійчук**

Доктор технічних наук, професор, завідувач кафедри

Кафедра «Комп'ютерні системи та мережі»  
Івано-Франківський національний технічний  
університет нафти і газу

вул. Карпатська, 15, м. Івано-Франківськ, 76019  
E-mail: gorb@nung.edu.ua

**М. А. Шуфнарівич**

Асистент

Кафедра «Медична інформатика, медична і  
біологічна фізика»

Івано-Франківський національний медичний  
університет

вул. Галицька, 2, м. Івано-Франківськ, 76019  
E-mail: shyfnarovich@gmail.com

### 1. Вступ

Газотранспортна мережа України включає близько 35 тис. км газопроводів, 71 компресорну станцію за-

гальною потужністю 5,4 млн. кВт. Важливою складовою газотранспортної мережі України є управління магістральних газопроводів "Прикарпаттрансгаз", яке в своєму складі має 18 компресорних станцій загаль-

ною продуктивністю 380×106 нм<sup>3</sup>/добу. Через річку Дністер також прокладені газопроводи такі, як: газопровід “Союз” (с.Коропець, Тернопільська область), газопровід “Прогрес” та газопровід “Урингой-Помари-Ужгород” (с.Михальче, Івано-Франківська область).

Тому, для забезпечення надійної та безаварійної роботи газопроводів, що проходять через річку Дністер, важливе значення має прогнозування його паводків. Часто паводки призводять до розмивання берегів, що може викликати пошкодження трубопроводу “Торжок-Долина” (Івано-Франківська область), який прокладено біля русла р. Дністер.

Згідно з результатами багаторічних спостережень за режимом р. Дністр на гідропості м. Галича, найхарактерніша риса водного режиму Дністра – дуже часті паводки протягом усього року як дощового, так і снігового походження. При цьому паводки високого рівня можуть бути в усі пори року. Загалом Дністер, як і всі Карпатські ріки, характеризується паводковим режимом. Висота паводків (біля Галича) коливається в межах 0,5-5 м.

Тому важливим питанням є побудова математичної моделі зміни рівня води ріки Дністер, що дасть можливість отримати точні результати прогнозування її стану та ввести запобіжні заходи для уникнення катастрофічних наслідків повеней.

## 2. Аналіз сучасних досліджень і публікацій

При дослідженні цілого ряду процесів, що відбуваються в екології мають справу з даними, які отримують у вигляді часових рядів. Крім того, часові ряди, які зв'язані з екологічними процесами часто мають помітні сезонні або періодичні компоненти; ці компоненти, як правило, змінюються у часі і у більшості випадків описуються сезонними статистичними моделями [1]. При цьому допускалось, що структура моделі певним чином вибрана і задачею ідентифікації була лише процедура визначення параметрів такої моделі, яка ґрунтувалась у більшості випадків на методі найменших квадратів.

Очевидно, що точність підгону моделі під емпіричні дані залежала від структури самої моделі. Наприклад, при виборі моделі у вигляді рівняння регресії при збільшенні кількості членів рівняння похибка, що визначається на всіх експериментальних точках монотонно падає. Як тільки число членів регресії стає рівним числу експериментальних точок, похибка стає рівною нулю. Відмітимо, що через задане число точок завжди можна провести нескінченно велике число кривих [2], тобто метод найменших квадратів породжує нескінченно велике число моделей для заданого числового ряду. Це твердження справедливе, якщо для вибору моделей користуватись лише одним критерієм. З теореми неповноти Геделя витікає, що користуючись даними навчальної послідовності, які були уже використані для знаходження коефіцієнтів за методом найменших квадратів, принципово неможливо знайти оптимальну і єдину модель. Тільки зовнішні доповнення дають змогу знайти єдину модель оптимальної складності [2].

Проблема створення математичних моделей складних техніко-екологічних процесів ґрунтовно розгля-

нута у цілому ряді праць А. Г. Івахненком та його учнів [1,2].

Із відомих методів прогнозування [1,2] найбільшої уваги заслуговує індуктивний метод самоорганізації складних моделей, який дає змогу в явному вигляді отримати емпіричну модель оптимальної складності. Недоліком методу є значні затрати машинного часу при практичній реалізації навіть для відносно нескладних моделей.

Метою роботи є розроблення методу побудови емпіричних моделей з використанням ідей генетичних алгоритмів для прогнозування екологічних процесів, який у значній мірі вільні від вказаного недоліку.

## 3. Основна частина

Протягом 2007 р. велись спостереження за рівнем води р. Дністер у районі с. Нижнів Івано-Франківської обл. з 1.04 по 31.08. За цей же період збирались дані про температуру повітря, кількість опадів, середню швидкість вітру та середньодобовий барометричний тиск.

В результаті проведених спостережень отримали множину експериментальних даних, що містить N = 184 точки.

Аналіз зміни рівня води у р. Дністер засвідчує, що з часом має місце тренд  $h(t)$ , який носить лінійний характер, та існує гармонічна складова  $G(t)$  зумовлена сезонною зміною метеорологічних умов [3], тобто

$$\tilde{H}_t = H_t + G(t) + h(t), \tag{1}$$

де  $\tilde{H}_t$  - поточний рівень води, см;

$G(t)$  - гармонічна складова рівня води, см;

$h(t)$  - лінійний тренд, см.

Із залежності (1), яка визначається зміною рівня води у р. Дністер, був виділений лінійний тренд

$$h(t) = \theta_0 + \theta_1 t, \tag{2}$$

де  $\theta_0, \theta_1$  - параметри лінійного тренду.

Коефіцієнти моделі (2) знайдено із застосуванням методу найменших квадратів. У результаті проведених обчислень, отримали такі значення коефіцієнтів лінійного тренду:  $\theta_0 = 308,5712$ ;  $\theta_1 = -0,4484$ .

Із числового ряду  $\tilde{H}_t$  виділяємо стаціонарну складову коливного процесу.  $G(t)$  подамо у вигляді гармонічного ряду [2] з некрatними частотами:

$$G(t) = A_0 + \sum_{j=1}^m (A_j \sin(\omega_j t) + B_j \cos(\omega_j t)), \tag{3}$$

де  $t$  - такти відліку часу,  $t = 1, 2, 3, \dots, N$ ;

$A_0, A_j, B_j$  - параметри гармонічного ряду (3);

$\omega_j = \omega_{j-1} + \Delta\omega_j$  - некрatні частоти,  $j = 1, 2, 3, \dots$

Для того, щоб за спостереженнями процесу  $G(t) = \tilde{H}_t - h(t)$ , можна було б оцінити параметри ряду (3), необхідне виконання умови [2]  $N \geq 3m + 1$ .

Суму декількох гармонік ряду (3), у якому коефіцієнти  $A_0, A_j, B_j$  визначені за методом найменших квадратів, а число гармонік і їх частоти вибрані так, щоб отримати мінімум деякого зовнішнього критерію

селекції називають [2] гармонічним трендом оптимальної складності.

Задача оптимального синтезу гармонічного ряду (3) розв'язана у роботі [3], що дозволило отримати наступну процедуру.

Виберемо функції

$$g(i+p) = \sum_{j=1}^m (A_j \sin((i+p)\omega_j) + B_j \cos((i+p)\omega_j)), \quad (4)$$

$$g(i-p) = \sum_{j=1}^m (A_j \sin((i-p)\omega_j) + B_j \cos((i-p)\omega_j)). \quad (5)$$

Для визначення частот  $\omega_j$  необхідно розв'язати рівняння

$$P_m z^m + P_{m-1} z^{m-1} + \dots + P_1 z + P_0 = 0, \quad (6)$$

де  $z = \cos \omega$ .

Рівняння (6) має  $m$  коренів, які однозначно визначають  $\omega_j$ ,  $j = 1, m$ .

Отже, для знаходження параметрів  $A_0$ ,  $A_j$ ,  $B_j$  і  $\omega_j$  гармонічного тренду спочатку потрібно визначити вагові коефіцієнти  $\alpha_p$ , з умови мінімізації нев'язки

$$B = \sum_{i=m+1}^{N-m} b_i^2, \quad (7)$$

де

$$b_i = g(i+m) - \sum_{p=0}^{m-1} \alpha_p (g(i+p) + g(i-p)) + g(i-m), \quad (8)$$

$$i = \overline{m+1, N-m}$$

характеризує точність, з якою коливний процес виражається через задану суму гармонічних складових. У формулі (8) величини  $g$  відповідних дискретних аргументів слід замінити на  $g_t = G(t)$ .

Отже, будемо розв'язувати задачу

$$\min_{\vec{\alpha}} J(\vec{\alpha}) = \sum_{i=m+1}^{N-m} \left( z_{i,m} - \sum_{p=0}^{m-1} \alpha_p g_{i,p} \right)^2, \quad (9)$$

де  $\vec{\alpha} = (\alpha_0, \alpha_1, \dots, \alpha_{m-1})^T$  - вектор вагових коефіцієнтів;

$$z_{i,m} = \overline{g(i+m)} + \overline{g(i-m)};$$

$$g_{i,p} = \overline{g(i+p)} + \overline{g(i-p)};$$

$T$  - символ транспонування матриць.

$$\min_{\vec{\alpha}} J(\vec{\alpha}) = (\vec{z}_m - F_m \vec{\alpha})^T (\vec{z}_m - F_m \vec{\alpha}), \quad (10)$$

$$\text{де } \vec{z}_m = \begin{bmatrix} \overline{g(2m+1)} + \overline{g(1)} \\ \overline{g(2m+2)} + \overline{g(2)} \\ \dots \\ \overline{g(N)} + \overline{g(N-2m)} \end{bmatrix};$$

$$F_m = \begin{bmatrix} 2\overline{g(m+1)} & \overline{g(m+2)} + \overline{g(m)} & \overline{g(m+3)} + \overline{g(m-1)} & \dots & \overline{g(2m)} + \overline{g(2)} \\ 2\overline{g(m+2)} & \overline{g(m+3)} + \overline{g(m+1)} & \overline{g(m+4)} + \overline{g(m)} & \dots & \overline{g(2m+1)} + \overline{g(3)} \\ \dots & \dots & \dots & \dots & \dots \\ 2\overline{g(N-m)} & \overline{g(N-m+1)} + \overline{g(N-m-1)} & \overline{g(N-m+2)} + \overline{g(N-m-2)} & \dots & \overline{g(N-1)} + \overline{g(N-2m+1)} \end{bmatrix}.$$

Мінімізація виразу (10) приводить до нормального рівняння Гауса, яке у матричній формі матиме вигляд:

$$F_m^T F_m \vec{\alpha} = F_m^T \vec{z}_m. \quad (11)$$

Із останнього рівняння можна знайти

$$\vec{\alpha} = (F_m^T F_m)^{-1} F_m^T \vec{z}_m. \quad (12)$$

Використовувати дану формулу (12) можна лише тоді, коли розмірність вектора  $\vec{\alpha}$  невелика і матриця  $F_m^T F_m$  є добре обумовленою. Якщо умова не виконується, то для знаходження  $\vec{\alpha}$  слід розв'язувати рівняння (11) одним із числових методів, наприклад, методом Гауса зі зворотнім ходом [4].

Нами запропонований інший підхід побудови математичних моделей коливних процесів, який базується на ідеях генетичних алгоритмів. Суть, якого у наступному.

Вся множина точок спостережень за вихідною величиною процесу або явища розбивається на три частини у такій пропорції [5]:  $N_R = 0,7N$ ,  $N_Q = 0,2N$  і  $N_S = 0,1N$ . Для множини даних  $N_R + N_Q$  визначаються вагові коефіцієнти  $\alpha_p$  як розв'язок лінійного алгебраїчного рівняння (11) за методом виключення Гауса з вибором головного елемента [4]. Розв'язок рівняння (6) відносно змінної  $z$  дає можливість знайти частоти  $\omega_j$ ,  $j = 1, m$ . За відомими частотами  $\omega_j$  на множині точок  $N_R + N_Q$  необхідно знайти параметри моделі (3)  $A_0$ ,  $A_j$  і  $B_j$ .

Поставлену задачу будемо розв'язувати, використовуючи генетичні алгоритми [6]. Утворимо хромосоми довжиною  $m$ , в якій на  $i$ -тому місці буде стояти нуль або одиниця в залежності від того чи частота  $\omega_j$  вилучена із вибраного повного ряду  $m$  чи залишена. Набір хромосом утворює популяцію. Із всієї популяції вибирають особі, що найбільше пристосовані, тобто такі, які мають найбільше (найменше) значення функції пристосованості. У задачі синтезу моделей коливних процесів функцією пристосованості виступає комбінований критерій селекції [7]

$$\rho = \sqrt{n_d^2 + B^2}, \quad (13)$$

де  $n_d^2$  - критерій зміщення, який обчислюється за такою формулою:

$$n_d^2 = \frac{\sum_{i=1}^N (g_i(R) - g_i(S))^2}{\sum_{i=1}^{N-2} g_i}, \quad (14)$$

де  $B$  - функція нев'язки, що визначається як (7);

$g_i(R)$ ,  $g_i(S)$  - величини, значення яких обчислені відповідно на множині точок  $N$  за формулою (3), а коефіцієнти моделі (3) знайдені відповідно на множинах  $N_R + N_Q$  і  $N_S$ .

Отже, задача синтезу моделі коливного процесу зводиться до вибору із початкової популяції такої хромосоми, яка забезпечує мінімальне значення критерію селекції (14).

Генетичний алгоритм складається із наступних кроків [6].

*К1. Формування початкової популяції (ініціалізація).*

На першому кроці роботи алгоритму випадковим чином формується популяція із  $I$  осіб, кожна із яких є хромосоною довжиною  $m$ .

*К2. Оцінка пристосованості хромосоми у популяції.*

Для кожної хромосоми обчислюється критерій селекції (14). А саме, у відповідності з моделлю (3) формується матриця

$$F = \begin{bmatrix} 1 & \sin \omega_1 & \cos \omega_1 & \sin \omega_2 & \cos \omega_2 & \dots & \sin \omega_m & \cos \omega_m \\ 1 & \sin(2\omega_1) & \cos(2\omega_1) & \sin(2\omega_2) & \cos(2\omega_2) & \dots & \sin(2\omega_m) & \cos(2\omega_m) \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & \sin(N\omega_1) & \cos(N\omega_1) & \sin(N\omega_2) & \cos(N\omega_2) & \dots & \sin(N\omega_m) & \cos(N\omega_m) \end{bmatrix}$$

Оскільки, кожній частоті  $\omega_j$  відповідає пара коефіцієнтів  $A_j, B_j$ , то у сформованій хромосомі подвоюємо одиниці і нулі. Якщо, наприклад, згенерована на першому кроці хромосома була такою:  $Ch = [1001011]$ , то після подвоєння вона набуде вигляду:  $Ch_d = [11000011001111]$ . У моделі (3) завжди присутній коефіцієнт  $A_0$ , то до хромосоми  $Ch_d$  на першу позицію додаємо одиничний ген і отримуємо  $Ch_{d0} = [111000011001111]$ .

У відповідності до отриманої хромосоми  $Ch_{d0}$  із матриці  $F$  формуємо нову матрицю  $F_{new}$ , шляхом вилучення тих стовпців із матриці  $F$ , які асоційовані із нулями хромосоми  $Ch_{d0}$ . Із отриманої матриці утворимо дві матриці  $F_R$  і  $F_S$ , розмірами  $(N_R + N_Q) \times m_1$  і  $N_S \times m_1$ .

На множинах точок  $N_R + N_Q$  і  $N_S$  обчислюються ненульові коефіцієнти  $A_0, A_j$  і  $B_j$  моделі (3), шляхом розв'язку нормального рівняння Гауса

$$F_R^T F_R \bar{A}_R = F_R^T \bar{g}_R, \tag{15}$$

$$F_S^T F_S \bar{A}_S = F_S^T \bar{g}_S, \tag{16}$$

де  $\bar{A}_R = (A_0^{(R)}, A_1^{(R)}, B_1^{(R)}, A_2^{(R)}, B_2^{(R)}, \dots, A_{m_1}^{(R)}, B_{m_1}^{(R)})$ ,

$\bar{A}_S = (A_0^{(S)}, A_1^{(S)}, B_1^{(S)}, A_2^{(S)}, B_2^{(S)}, \dots, A_{m_1}^{(S)}, B_{m_1}^{(S)})$  - вектори параметрів моделі, яка асоційована з черговою хромосоною із початкової популяції і обчислені за формулами (13) і (14);

$\bar{g}_R = (g^{(-1)}, g^{(-2)}, \dots, g^{-(N_R+N_Q)})^T$ ,  $\bar{g}_S = (g^{(-1)}, g^{(-2)}, \dots, g^{-(N_S)})^T$  - вектори експериментальних даних на множині точок  $N_R + N_Q$  і  $N_S$ .

За відомою сукупністю коефіцієнтів  $\bar{A}_R$  і  $\bar{A}_S$  моделі (3) на множині точок  $N$  обчислюють

$$g(R) = F_{new} \bar{A}_R, \tag{17}$$

$$g(S) = F_{new} \bar{A}_S. \tag{18}$$

За формулою (13) обчислюють критерій селекції, де  $B$  знаходять у відповідності з (7) і (8). Значення критерію селекції обчислюють для кожної хромосоми і в результаті отримують множину значень  $\rho_i, i=1, M$ , де  $M$  - кількість хромосом у популяції.

*К3. Перевірка умови зупинки алгоритму.*  
Визначають

$$\rho_{min} = \min_{i \in M} \rho_i. \tag{19}$$

Якщо мінімальне значення (19) критерію селекції (13) не перевершує деякого додатного значення  $\epsilon$ , то відбувається зупинка алгоритму. Зупинка алгоритму, також, може відбутися у випадку, коли його виконання не приводить до покращення функції пристосування або у тому випадку, коли алгоритмом уже виконано задане число ітерацій.

Після виконання однієї із трьох умов із популяції вибирається хромосома  $ch^*$ , для якої виконується умова (19). Після операції подвоєння і приєднання одиничного гену до хромосоми  $Ch_d^*$  отримуємо -  $Ch_{d0}^*$ . Ця хромосома задає структуру моделі оптимальної складності і формує матрицю  $F^*$  таким чином, що із початкової матриці  $F$  вилучаються стовпці, які асоційовані з нульовими генами хромосоми  $Ch_{d0}^*$ . Перерахунок параметрів моделі (3) здійснюється на множині всіх точок початкового масиву даних.

*К4. Селекція хромосом.*

За розрахованими на другому кроці значеннями функції пристосування здійснюється відбір тих хромосом, які будуть брати участь в створенні потомків для наступної популяції. Найбільші шанси у створенні нової популяції мають хромосоми з найкращим значенням функції пристосування, тобто такі, що забезпечують мінімальне значення критерію селекції (13).

Найбільш поширеними методами селекції [6] є метод рулетки і метод турнірної селекції. У даному алгоритмі використано турнірний метод. При турнірній селекції всі хромосоми розбиваються на підгрупи з наступним вибором із кожної утвореної підгрупи хромосоми з найкращою пристосованістю. Підгрупи можуть мати довільний розмір, але частіше за все популяції ділять на підгрупи по 2 – 3 особи у кожній.

*К5. Формування нової популяції потомків.*

Популяції потомків формується за допомогою двох основних операторів - схрещування і мутації. Слід відмітити, що у генетичному алгоритмі схрещування проводиться майже завжди, а мутація – досить рідко. Вірогідність схрещування досить велика ( $0,5 \leq P_c \leq 1$ ), тоді як ймовірність мутації вибирається досить малою ( $0 \leq P_m \leq 0,1$ ).

Оператор мутації з ймовірністю  $P_m$  змінює значення гена в хромосомі на протилежне, тобто з 1 на 0 чи з 0 на 1. Ймовірність мутації  $P_m$  може здійснюватися випадковим вибором числа із інтервалу  $[0;1]$  для кожного гена і відбором для виконання цієї операції тих генів, для яких розігране число виявиться меншим або рівним  $P_m$ . Мутація може здійснюватись як над пулом родичів, так і над пулом потомків.

Оператор схрещування складається із двох етапів. На першому етапі формуються підгрупи із  $z$  осіб, звідки вибирається найкраща хромосома за критерієм селекції  $R(\text{Ch}_{d0}^*) = \min R(\text{Ch}_{d0,i})$ . У результаті отримуємо нову популяцію хромосом, до якої застосовують оператор другого етапу.

На другому етапі здійснюється схрещування. Для цього із пулу родичів  $M(k)$  випадковим чином з ймовірністю  $P_c$  утворюють пари у такий спосіб. Із популяції осіб випадковим чином вибирається пара хромосом. Генерується випадкове число  $P_z$  із інтервалу  $[0; 1]$  і якщо його значення не більше ніж  $P_c$ , то над парою хромосом здійснюється схрещування. У протилежному випадку пара хромосом залишається без зміни. Потім для кожної пари родичів розігрується позиція гена (локус) в хромосомі, яка визначає точку схрещування. Якщо хромосома кожного із родичів включає у себе  $m$  генів, то точка схрещування  $L_c$  - це натуральне число, яке менше  $m$ . Тому фіксація точки схрещування зводиться до випадкового вибору цілого числа із інтервалу  $[1; L_c - 1]$ .

Дія оператора схрещування приводить до того, що із пари родичів утворюється нова пара потомків наступним чином: перший потомок у парі, хромосома, якого на позиціях від 1 до  $L_c$  складається із ген першого родича, а на позиції від  $L_c + 1$  до  $m$  із ген другого родича; другий потомок у парі, хромосома, якого на позиціях від 1 до  $L_c$  складається із ген другого родича, а на позиції від  $L_c + 1$  до  $m$  із ген першого родича.

Після виконання оператора схрещування відбувається перехід до K2.

З використанням розробленого методу у середовищі MatLab написана програма виділення гармонічного тренду  $G(t)$  з некротними частотами. Було вибрано максимальне число частот  $m = 30$ ; число точок спостережень  $N = 160$ .

Ймовірність схрещування  $P_c = 0,9$ , а ймовірність мутації складала  $P_m = 0,1$ . Таким чином, максимальне число коефіцієнтів моделі (3), які визначались, складало  $2m + 1 = 61$ ; із них - 8 нульові. Результат роботи програми відтворює рис. 1, де знаком «o» відмічені експериментальні дані, а «+» - результат розрахунку за формулою (3).

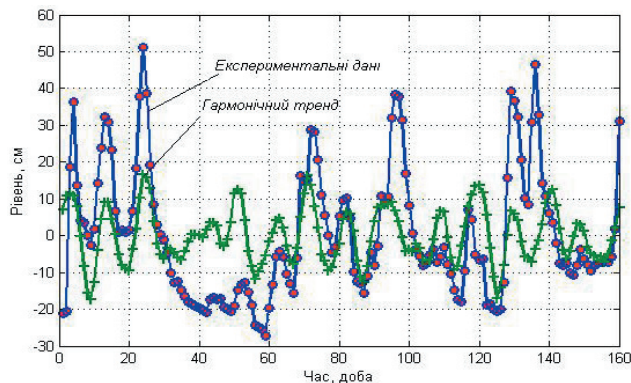


Рис. 1. Гармонічний тренд коливного процесу (р. Дністер)

Після виділення із експериментальних даних гармонічного тренду отримують залишок (рис. 2). Величину цього залишку визначимо із рівняння (1) за умови, що попередньо визначені лінійний тренд і

гармонічна складова -  $H_t = \tilde{H}_t - (G(t) + h(t))$ . Величина  $H_t$  є функцією параметрів, що визначають погодні умови у районі спостережень, тобто

$$H_t = \phi(T_t, f_{t-k}, v_t, p_t), \quad (20)$$

де  $T_t$  - середньодобова температура повітря, °C;  
 $f_t$  - кількість опадів, мм/добу;  
 $t$  - поточний дискретний час, доби;  
 $k$  - зсув у часі;  
 $v_t$  - середньодобова швидкість вітру, м/с;  
 $p_t$  - середньодобовий барометричний тиск, мм. рт. ст.

Розглянемо даний процес як систему, функціонування якої характеризується сукупністю вхідних величин  $x_1 = T_t, x_2 = f_t, x_3 = f_{t-1}, x_4 = f_{t-2}, x_5 = f_{t-3}, x_6 = v_t, x_7 = p_t$  і вихідною величиною  $y = H_t$ .

Значення величини  $Y_t = H_t$  визначені у дискретні моменти часу  $t = 1, \dots, N$ . Вхідні величини  $x_j, j = 1, k$  у кожному спостереженні  $t$  приймають певне значення так, що їх сукупність утворює матрицю

$$X = \begin{bmatrix} x_1^{(1)} & x_2^{(1)} & \dots & x_k^{(1)} \\ x_1^{(2)} & x_2^{(2)} & \dots & x_k^{(2)} \\ \dots & \dots & \dots & \dots \\ x_1^{(N)} & x_2^{(N)} & \dots & x_k^{(N)} \end{bmatrix}.$$

Співвідношення (20) будемо шукати у вигляді полінома

$$Y_t = \sum_{i=0}^{M-1} a_i \prod_{j=1}^k (x_j^{(t)})^{s_{ji}}, \quad t = \overline{1, N}. \quad (21)$$

де  $M$  - кількість членів полінома;  
 $a_i$  - коефіцієнти полінома;  
 $s_{ji}$  - степені аргументів, які повинні задовольняти обмеженню  $\sum_{j=1}^k s_{ji} \leq m$ .

Число членів  $M$  полінома (21) визначають за такою формулою [8]:

$$M = \frac{(m+k)!}{m!k!}. \quad (22)$$

Систему рівнянь (21) зручно подати у матрично-векторній формі

$$\bar{y} = F \bar{a}, \quad (23)$$

де  $\bar{y} = (y_1, y_2, \dots, y_N)^T$  - обчислене значення виходу моделі (26) у кожній точці спостережень;

$F$  - матриця розміром  $N \times M$ , елементи якої добутки аргументів при параметрах  $a_i$ , тобто

$$F = \begin{bmatrix} \prod_{j=1}^k (x_j^{(1)})^{s_{j0}} & \prod_{j=1}^k (x_j^{(1)})^{s_{j1}} & \dots & \prod_{j=1}^k (x_j^{(1)})^{s_{jM-1}} \\ \prod_{j=1}^k (x_j^{(2)})^{s_{j0}} & \prod_{j=1}^k (x_j^{(2)})^{s_{j1}} & \dots & \prod_{j=1}^k (x_j^{(2)})^{s_{jM-1}} \\ \dots & \dots & \dots & \dots \\ \prod_{j=1}^k (x_j^{(N)})^{s_{j0}} & \prod_{j=1}^k (x_j^{(N)})^{s_{j1}} & \dots & \prod_{j=1}^k (x_j^{(N)})^{s_{jM-1}} \end{bmatrix};$$

$\bar{a} = (a_0, a_1, \dots, a_{M-1})^T$  - вектор параметрів моделі (21).

На практиці, як правило, структура моделі (21) невідома, що приводить до необхідності довільного вибору як числа функцій, так і вигляду самих функцій у моделі (21). Тому для вибору структури моделі (21) був запропонований індуктивний метод самоорганізації моделей [2], ідейну сторону якого визначає теорема Геделя. Стосовно задачі визначення структури моделі (21) геделівський підхід означає застосування зовнішнього критерію, який дає можливість однозначного вибору єдиної моделі із заданого класу моделей. Визначення зовнішнього критерію здійснюється на нових даних, які не використовувались при синтезі моделі (21). Отже, всі дані, які отримані у результаті експерименту, розбиваються на дві частини навчальну -  $N_R$  і перевіірочну -  $N_Q$ .

У більшості випадків для вибору структури моделі використовують критерії регулярності

$$\Delta^2(Q) = \frac{\sum_{i=1}^{N_Q} (Y_i - y_i)^2}{\sum_{i=1}^{N_Q} Y_i^2} \tag{24}$$

і мінімуму зміщення

$$\Delta^2(R, Q) = \frac{\sum_{i=1}^N (y_i(R) - y_i(Q))^2}{\sum_{i=1}^N Y_i^2} \tag{25}$$

Якщо вибраний критерій регулярності (24), то вибирають наступний розподіл даних експерименту [5]:  $N_R = 0,7N$  і  $N_Q = 0,3N$ , а при виборі критерію мінімуму зміщення (25) -  $N_R = 0,5N$  і  $N_Q = 0,5N$ .

Реалізація індуктивного методу самоорганізації моделей здійснюється поетапно: перший етап – генерація моделей-претендентів (у певному порядку підвищення складності); другий етап – відбір найкращої моделі за критерієм селекції (24) або (25).

Проаналізувавши недоліки відомих методів генерації моделей-претендентів, вибрано комбінаторний метод, оскільки на відміну від інших методів він дає можливість отримати модель, де аргументами виступають вхідні величини системи. Зауважимо, при побудові моделі комбінаторним методом, прийшлося би перебрати  $2^M - 1$  моделей. У відповідності з формулою (22) для  $n = 7$  і  $m = 4$  -  $M = 330$ . Тоді  $2^M - 1 = 2,1873 \cdot 10^{99}$  варіантів, що практично неможливо реалізувати за допомогою сучасних персональних комп'ютерів. Для зняття проблеми великої розмірності застосуємо генетичний підхід.

При комбінаторному методі синтезу моделі із повного полінома (21) отримують емпіричну модель, де частина параметрів приймає значення нуль. Інші параметри, що залишились, будуть відмінні від нуля. Утворимо упорядковану структуру довжиною  $M$ , в якій на  $i$ -тому місці буде стояти одиниця або нуль в залежності від того чи параметр  $a_i$ ,  $i = 1, M$  моделі (21) відмінний від нуля, чи нульовий.

Таким чином, задача синтезу емпіричної моделі полягає у виборі із початкової популяції такої, хромосоми, яка забезпечує найкраще значення функції пристосування (мінімальне значення критерію селекції

(24) або (25)). Алгоритм розв'язку поставленої задачі аналогічний раніше розробленому для виділення гармонічного тренду.

На основі розробленого алгоритму була написана програма у середовищі MatLab для побудови математичної моделі залишку, який отримали після вилучення лінійного тренду і гармонічної складової. Було вибрано  $m = 4$ . З використанням розробленої програми синтезована модель, яка вміщує 154 ненульових і  $330 - 154 = 176$  нульових параметрів  $a_i$ ,  $i = 0, M - 1$  полінома (21).

Результати роботи програми відтворює рис. 2, де через «o» позначені експериментальні дані, а через «+» - значення  $y$ , які обчислені як вихід синтезованої моделі.

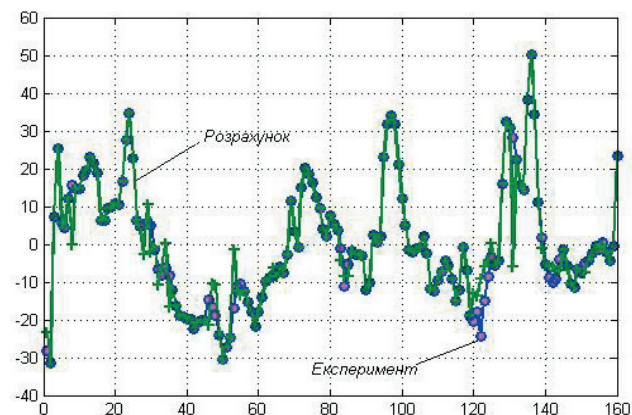


Рис. 2. Залежності рівня води від параметрів погодних умов (після вилучення лінійного і гармонічного трендів)

Адекватність моделі перевірялась за допомогою коефіцієнта кореляції  $K_{Yy}$  між значеннями  $Y_i$  та її виходом  $y^{(i)}$

$$K_{Yy} = \frac{\sum_{i=1}^N (Y_i - \bar{Y})(y^{(i)} - \bar{y})}{\sqrt{\sum_{i=1}^N (Y_i - \bar{Y})^2 \sum_{i=1}^N (y^{(i)} - \bar{y})^2}} \tag{26}$$

де  $\bar{Y} = \frac{1}{N} \sum_{i=1}^N Y_i$ ,  $\bar{y} = \frac{1}{N} \sum_{i=1}^N y^{(i)}$  - оцінки математичних сподівань для величин  $Y_i$  і  $y^{(i)}$ .

Було отримано:  $K_{Yy} = 0,9662$ , що свідчить про високу степінь кореляції між величинами  $Y_i$  і  $y^{(i)}$ .

Знайдені залежності  $h(t)$ ,  $G(t)$  і  $u_t$  дають можливість знайти

$$\tilde{N}_t = G(t) + h(t) + u_t, \tag{27}$$

де  $h(t)$ ,  $G(t)$  і  $u_t$  - обчислювались відповідно за формулами (2), (3) і (21).

Графік залежності (27) показаний на рис. 3, де «+» позначені обчислені значення за формулою (27), а значком «o» відмічені експериментальні значення рівня води у р. Дністер.

Із графіка видно, що мають місце досить задовільні збіги між розрахунковими і експериментальними даними.

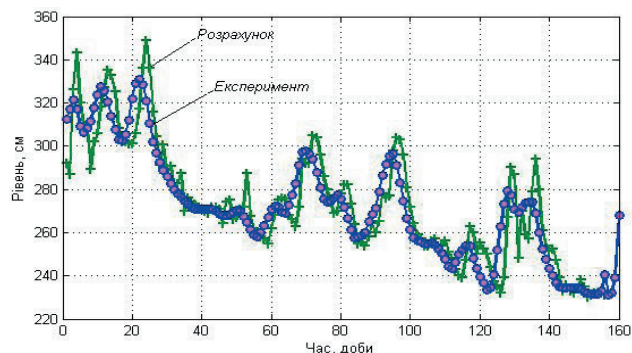


Рис. 3. Залежність рівня води у р. Дністер від погодних умов

Цінність виконаного прогнозу значною мірою визначається його якістю, яка залежить від ступеня збігу майбутнього значення процесу з оцінкою цього значення, зробленою заздалегідь за математичною моделлю.

Для оцінки якості прогнозу при самоорганізації прогнозуючих моделей з вхідних даних крім навчальної  $N_R$  і перевіркової  $N_Q$  вибірок виділялась частина даних на екзаменаційну вибірку  $N_S$ . На екзаменаційній множині  $N_S$  визначали середньоквадратичне відхилення обчисленого за допомогою моделі  $y_i$  (прогнозованого) значення і дійсного значення  $Y_i$  [10]:

$$\Delta^2(S) = \frac{\sum_{i=1}^{N_S} (Y_i - y_i)^2}{\sum_{i=1}^{N_S} Y_i^2} \rightarrow \min. \quad (27)$$

Написана програма обчислює значення критерію на основі екзаменаційної вибірки  $N_S = 24$ . Отримане значення  $\Delta^2(S) = 0,0058$  свідчить про досить непогану точність прогнозу рівня ріки Дністер.

На рис. 4 показані отримані на екзаменаційній множині експериментальні значення та значення за побудованою моделлю (28). З рисунка видно, що отримані значення за прогностичною моделлю (1) досить близькі до фактичних значень рівня води

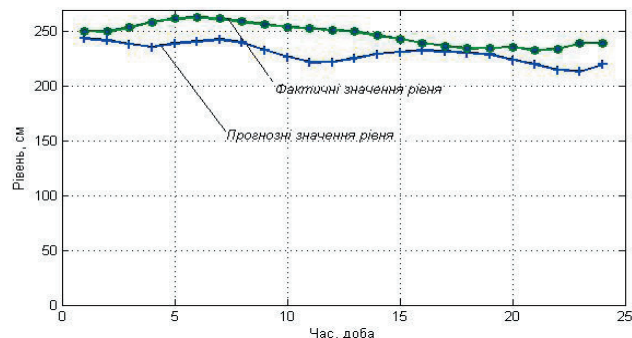


Рис. 4. Результати прогнозування рівня води на екзаменаційній множині

#### 4. Висновок

Таким чином, застосування ідей генетичних алгоритмів до побудови математичної моделі зміни рівня води у р. Дністер дало можливість отримати адекватну модель і значно зменшити об'єм обчислень. Перевірка точності прогнозу за отриманою емпіричною моделлю на екзаменаційній множині показала, що вона здатна з високою ймовірністю прогнозувати зміну рівня води у р. Дністер. Це свідчить про те, що отримана модель зміни рівня води у р. Дністер у залежності від погодних умов може бути використана при прогнозуванні повеней, що як показали події 2008 р. є досить актуальною задачею для Прикарпатського регіону.

#### Література

1. Бокс, Дж. Анализ временных рядов. Прогноз и управление [Текст] / Дж. Бокс, Г. Дженкинс. Пер. с англ. А. Л. Левшина. – М.: Мир, 1974. – 406 с.
2. Ивахненко, А. Г. Индуктивный метод самоорганизации моделей сложных систем: монография [Текст] / А. Г. Ивахненко. – К.: Наукова думка, 1981. – 296 с.
3. Горбійчук, М. І. Метод побудови математичних моделей складних процесів на засадах генетичних алгоритмів [Текст] / Горбійчук М. І., Шуфнарівич М. А. // Штучний інтелект. – 2010. - № 4. – С. 50 – 57.
4. Вержбицкий, В. М. Основы численных методов: учебник для вузов [Текст] / В. М. Вержбицкий. – М.: Высшая школа, 2002. – 840 с.
5. Ивахненко, А. Г. Справочник по типовым программам моделирования [Текст] / А. Г. Ивахненко, Ю. В. Коппа, В. С. Степашко и др.; под ред. А. Г. Ивахненко – К.: Техніка, 1980. – 180 с.
6. Рутковская, Д. Нейронные сети, генетические алгоритмы и нечеткие системы [Текст] / Д. Рутковская, М. Пилиньский, Л. Рутковский; пер. с польск. И. Д. Рудинского. – М.: Горячая линия-Телеком, 2004. – 452 с.
7. Ивахненко, А. Г. Помехоустойчивость моделирования: монография [Текст] / А. Г. Ивахненко, В. С. Степашко - Киев: Наук. думка, 1985. - 216 с.
8. Горбійчук, М. І. Индуктивный метод побудови математичних моделей газоперекачуваль-них агрегатів природного газу [Текст] / М. І. Горбійчук, М. І. Когутяк, Я. І. Заячук // Нафтова і газова промисловість. – 2008. - № 5. – С. 32 – 35.
9. Ермаков, С. М. Математическая теория оптимального эксперимента [Текст] / С. М. Ермаков, А. А. Жиглявский: учеб. пособие. – М.: Наука, 1987. – 320 с.
10. Ивахненко, А. Г. Самоорганизация прогнозирующих моделей [Текст] / А. Г. Ивахненко, Й. А. Мюллер – К.: Техніка, 1985; Берлин: ФЕБ Ферлаг Техник, 1984. – 223 с., ил. – Библиогр.: 219с.