

## Часовий аналіз даних гідрохімічного моніторингу питної води

АРХАНГЕЛЬСЬКА Ю.М.

Дніпропетровський національний університет

На підставі методу обробки, аналізу та прогнозування SSA-«Гусениця» і адаптивних методів прогнозування розроблено інформаційну технологію часового аналізу даних гідрохімічного моніторингу питної води.

На основани методу обробки, анализа и прогнозирования SSA-«Гусеница» и адаптивных методов прогнозирования разработана информационная технология временного анализа данных гидрохимического мониторинга питьевой воды.

On the basis of method of treatment, analysis and prognostication of SSA-«Gusenica» and adaptive methods of prognostication information technology of time analysis of data of the hydrochemical monitoring of drinking-water is developed.

**Постановка проблеми.** Дані гідрохімічного моніторингу питної води є основою для вивчення її хімічного забруднення. В свою чергу, хімічне забруднення, як об'єкт вивчення уявляє собою нестационарний динамічний ряд, тобто варіаційну послідовність разових концентрацій, що характеризує зміну рівнів дії в часі і в просторі [1]. В результаті такої структури при обробці даних гідрохімічного моніторингу питної води (ГХМ ПВ), виникає задача їх часового аналізу, що містить у собі вияв структури та просторо-часової зміни хімічного складу питної води.

**Аналіз останніх досліджень і публікацій, в яких започатковано розв'язання даної проблеми та постановка задачі.** Традиційно для вияву та оцінки структури часових рядів використовується швидке перетворення Фур'є, що дає достатньо прості для розрахунків формули розкладання досліджуваного тимчасового процесу  $f(t)$  в ряд по тригонометричних функціях однак в деяких випадках перетворення може дати невірні результати для сигналів з ділянками різкої зміни та ряди не завжди задовольняють вимозі періодичності і, як правило, задані на обмеженому відрізку часу [2]. Вирішення цих проблем може бути досягнуто за використанням методу обробки, аналізу та прогнозування часових рядів SSA-«Гусениця». Так у роботі [3] розглянуто приклади застосування методу при дослідженні періодичних компонент природних процесів на даних гідрохімічного моніторингу Чудсько-Псковського озера; оцінена тенденція збільшення площі криги Балтики за 273 – річним рядом спостережень та виявлено криві тренду у вигляді синусоїд з періодом біля 300 років. В процесі аналізу даних методом «Гусениця» було виділено періодичні компоненти в рядах озера Песьво та Удов'я.

Для оцінки поведінки даних гідрохімічного моніторингу використовуються методи прогнозування. Так проведення попереднього тренд-аналізу та прогнозування часових рядів гідрохімічних даних надано у роботах [4-

6]. Прогнозування гідрохімічних даних проводилось за методами: «Гусениця», гармонійних ваг, експоненційного прогнозування, імітаційним та адаптивними методами.

Аналізуючи представлені публікації, можна зробити висновок, щодо використання цих методів для вирішення окремих задач. Комплексне ж застосування цих математичних апаратів дозволить провести більш адекватну оцінку часової зміни хімічного складу питної води.

В зв'язку з цим постає задача в необхідності розробки інформаційної технології часового аналізу даних ГХМ ПВ.

**Виклад основного матеріалу дослідження.** Часовий аналіз за даними ГХМ ПВ було проведено за 11 точками водозабору питної води:

- лівобережна частина Дніпропетровського району: № 1,2,3 с. Чумаки, с.м.т. Ювілейне, № 1,2 с. Партизанське, с. В.Маївка, с. Степне, с. Заря;
- правобережна частина Дніпропетровського району: с. НовоМиколаївка, с. Сурсько – Литовське за гідрохімічними показниками:  $Fe$ ,  $F$ ,  $Cl$ ,  $Cu$ ,  $SO_4^{2-}$ ,  $NH_4^+$ ,  $NO_3$ ,  $NO_2$ , сухий залишок, загальна жорсткість на період з 2002-2008рр.

Дані о концентрації гідрохімічних показників подано у вигляді масиву:

$$\{g(t), t = \overline{1, M}\}, \quad (1)$$

де  $M$  - момент часу дослідження;  $g(t)$  - концентрація гідрохімічного показника в  $t$ -й момент часу.

Обчислювальний процес часового аналізу даних ГХМ ПВ (1) складається з наступних етапів:

1. Етап аналізу структури ряду за даними ГХМ ПВ .
2. Етап прогнозування за даними ГХМ ПВ.
3. Етап вибору адекватного методу прогнозування за даними ГХМ ПВ.

Схема інформаційної технології часового аналізу за даними ГХМ ПВ представлено на рис. 1.

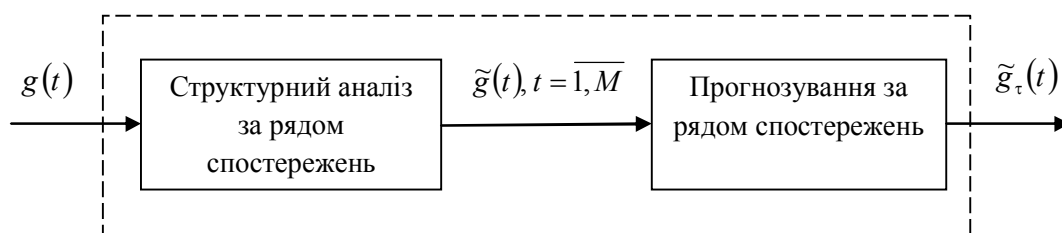


Рис. 1. Інформаційна технологія часового аналізу за даними ГХМ ПВ.

Згідно [3, 7], для подальшого викладення матеріалу введемо наступні поняття:

$$G = (G_1, G_2, \dots, G_K) = \begin{pmatrix} g_1 & g_2 & g_3 & \dots & g_L \\ g_2 & g_3 & g_4 & \dots & g_{L+1} \\ g_3 & g_4 & g_5 & \dots & g_{L+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ g_K & g_{K+1} & g_{K+2} & \dots & g_M \end{pmatrix} \quad (2)$$

де  $G$  - траєкторна матриця ряду  $g(t)$ ,  $K = M - L + 1$ ,  $1 < L < M$ .

$$g_{ij}^* = \frac{(g_{ij} - \mu_j)}{\delta_j}, \quad (3)$$

де  $g_{ij}^*$  - перетворення елементів матриці (1),  $i = \overline{1, K}$ ,  $j = \overline{1, L}$ .

$$\mu_j = \begin{cases} 0 \\ \bar{g}_j \end{cases}, \quad \bar{g}_j = \frac{1}{K} \sum_{i=1}^K g_{ij}, \quad (4)$$

$$\delta_j = \begin{cases} 1 \\ s_j \end{cases}, \quad s_j = \sqrt{\frac{1}{K} \sum_{i=1}^K (g_{ij} - \bar{g}_j)^2}, \quad (5)$$

$$S = \frac{1}{K} G^* G^{*T} \quad (6)$$

де  $S$  - матриця других моментів за траєкторною матрицею  $G^*$ .

$$S = V \Lambda V^T, \quad (7)$$

де

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & 0 & \dots & 0 \\ 0 & 0 & \dots & \lambda_L \end{pmatrix}, \quad (8)$$

де  $\Lambda$  - діагональна матриця власних чисел матриці других моментів  $S$ .

Ядро інформаційної технології складають: метод обробки, аналізу та прогнозування SSA - «Гусениця» та даптивні методи прогнозування з сезонним ефектом та без сезонної складової.

$$V = (v_1, v_2, \dots, v_L) = \begin{pmatrix} v_{11} & v_{12} & \dots & v_{1L} \\ v_{21} & v_{22} & \dots & v_{2L} \\ \dots & \dots & \dots & \dots \\ v_{L1} & v_{L2} & \dots & v_{LL} \end{pmatrix} \quad (9)$$

де  $V$  - матриця власних векторів матриці других моментів  $S$ .

$$U_{G^*} = V^T G^* = (U_1 : \dots : U_L)^T \quad (10)$$

де  $U_{G^*}$  - операція отримання нецентрального головного компонента.

$$\tilde{G}^* = (v_1, v_2, \dots, v_L) \begin{pmatrix} U_1^T \\ \dots \\ U_L^T \end{pmatrix} = \sum_{i=1}^{L^*} v_i U_i^T, \quad (11)$$

де  $\tilde{G}^*$  - відтворення  $g(t)$  за першими  $L^*$  головними компонентами.

$$\tilde{g}_s = \begin{cases} \frac{1}{s} \sum_{i=1}^s \tilde{g}_{s-i+1,i} & \text{для } 1 \leq s \leq L^*, \\ \frac{1}{L^*} \sum_{i=1}^{L^*} \tilde{g}_{s-i+1,i} & \text{для } L \leq s < K^*, \\ \frac{1}{M-s+1} \sum_{i=1}^{M-s+1} \tilde{g}_{K^*-i+1, s+i-K^*} & \text{для } K^* \leq s < M. \end{cases} \quad (12)$$

де  $\tilde{g}_s$  - діагональне усереднювання матриці  $\tilde{G}$ .

$$\begin{cases} \sum_{j=1}^{L^*} h_j v_{1j} = \tilde{g}_{M-L^*+k+2} \\ \dots \\ \sum_{j=1}^{L^*} h_j v_{1L^*-1} = \tilde{g}_{M+k-1} \end{cases} \quad (13)$$

Таким чином, з урахуванням понять (2) - (13), алгоритм структурного аналізу за рядом гідрохімічних показників (1) подано нижче.

**Алгоритм 1.**

1. Задається параметр методу «довжина гусениці»  $L$ ,  $1 < L < M$ .
2. Проводиться розгортка початкового одновимірного ряду (1) у послідовність  $L$ -вимірних векторів, число яких дорівнює  $K = M - L + 1$ , у вигляді (2).
3. Обчислюється матриця  $G^*$ , елементи якої отримують за (3) з урахуванням (4) та (5).
4. Обчислюється матриця (6).
5. Проводиться сингулярне розкладання матриці (6) у вигляді (7) з урахуванням (8) та (9).
6. Обчислюються головні компоненти згідно (10).
7. Задається кількість головних компонент  $L^*$ ,  $L^* \leq L$ .
8. Проводиться відтворення багатовимірного ряду згідно (11) з урахуванням (10).
9. Якщо  $\mu_j = \bar{g}_j$ , проводиться децентрування матриці  $\tilde{G}^*$ , результатом якої є матриця  $\tilde{G}$ .
10. Якщо  $\delta_j = s_j$ , проводиться денормування матриці  $\tilde{G}^*$ , результатом якої є матриця  $\tilde{G}$ .
11. Виконується ганкелізація матриця  $\tilde{G}$  згідно (12).

Застосування методу SSA -«Гусениця» дозволить виявлення та видалення неперіодичних коливань за даними ГХМ ПВ.

Наступним етапом часового аналізу є прогнозування за даними ГХМ ПВ за використанням адаптивних методів прогнозування з сезонним ефектом та без сезонної складової та методу SSA -«Гусениця».

**Теорема 1:** Нехай ряд (1) припускає продовження  $\{g(t), t = M + 1\}$ . Тоді він може бути продовжений необмежено на будь-яке число кроків, причому це продовження єдине [3].

Для проведення процесу прогнозування ряду спостережень гідрохімічних показників (1) на  $\tau$  кроків, тобто отримання продовження ряду у вигляді  $\{\tilde{g}_\tau(t) \mid t = \overline{1, M}; \tau = \overline{1, 10}\}$  з урахуванням теореми 1, представлено алгоритм 2.

**Алгоритм 2.**

1. Виконуються пункти 1- 4 алгоритму 1.
2. Вводиться  $\tau$  - кількість кроків прогнозування.
3. Знаходиться базис  $v_1, v_2, \dots, v_L$ , що відповідає не нульовим власним числам матриці (6).
4.  $\tau' = 1$ ;
5. Система рівнянь (13) перевіряється на сумісність при  $\tau' = 1$ .
6. Якщо система рівнянь (13) не сумісна, тоді ряд (1) не має продовження.
7. Якщо система рівнянь (13) має розв'язок, тоді знаходимо продовження ряду  $\tilde{g}_\tau(t)$  згідно (13)
8. Збільшується значення  $\tau' = \tau' + 1$ .
9. Якщо  $\tau' \leq \tau$  повторюються пункти (6)-(8).
10. Закінчити роботу алгоритму.

Основною адаптивних методів є модель. Первинна оцінка параметрів цієї моделі ґрунтується на даних базового (початкового) тимчасового ряду. На основі нових даних, що отримуються на кожному наступному кроці, відбувається коректування параметрів моделі в часі, їх адаптація до нових умов розвитку явища, що безперервно змінюються. Таким чином, модель постійно «вбирає» нову інформацію і пристосовується до неї [8].

В якості адаптивних моделей прогнозування виступають моделі з сезонним ефектом та без сезонної складової.

В основі адаптивних моделей без сезонної складової лежить гіпотеза про те, що прогноз може бути отримано з рівняння

$$\tilde{g}_\tau(t) = \hat{a}_{0,t} + \hat{a}_{1,t}\tau, \quad (15)$$

де  $\hat{a}_{0,t}$  - поточне значення після виключення сезонності;  $\hat{a}_{1,t}$  - коефіцієнт лінійного тренду;  $\tau$  - кількість кроків прогнозування.

Процедури обчислення параметрів  $\hat{a}_{0,t}, \hat{a}_{1,t}$  та похибки прогнозу відповідно до моделі згідно [7] подано нижче.

**Процедура 1. Модель Тейла-Вейджа.**

$$\begin{aligned} \hat{a}_{0,t} &= \hat{a}_{0,t-1} + \hat{a}_{1,t-1} + \alpha_1 e_1(t-1), \\ \hat{a}_{1,t} &= \hat{a}_{1,t-1} + \alpha_2 e_1(t-1), \end{aligned} \quad (16)$$

$$e_1(t) = g(t) - \hat{a}_{0,t} - \hat{a}_{1,t} = (a_{0,t} - \hat{a}_{0,t}) + (a_{1,t} - \hat{a}_{1,t}) + v_{t+1} + \varepsilon_{t+1}$$

де  $v_t, \varepsilon_t$  - похибки, що визначаються відповідно «приростом» прогнозу і погрішністю, які мають нормальний закон розподілу  $N_1(0, \sigma_v), N_2(0, \sigma_\varepsilon)$ ;  $0 < \alpha_1, \alpha_2, \alpha_3 < 1$  - параметри згладжування, причому  $\alpha_2 \approx 0$ .

**Процедура 2. Модель Дж.Бокса, Ч.Дженкінса.**

$$\begin{aligned} \hat{a}_{0,t} &= \alpha_1 g(t) + (1 - \alpha_1)(\hat{a}_{0,t-1} + \hat{a}_{1,t-1}) + \alpha_3 (e(t) - e(t-1)), \\ \hat{a}_{1,t} &= \alpha_2 (\hat{a}_{0,t-1} - \hat{a}_{1,t-1}) + (1 - \alpha_2) \hat{a}_{1,t-1} \end{aligned} \quad (17)$$

Оптимальні значення  $\alpha_1, \alpha_2$  знаходяться із умови мінімуму середнього квадрату похибки:

$$\begin{aligned} \alpha_1 &= \frac{2h}{1+h}, \quad \alpha_2 = h, \quad g^2 = \frac{\sigma_v^2}{\sigma_\varepsilon^2}, \\ h &= \sqrt{-0,125g^2 + 0,5g\sqrt{1+0,0625g^2}}, \end{aligned} \quad (18)$$

$$\sigma_\varepsilon^2 = \frac{1-h}{1+h} D_e(1),$$

$$D_e(t) = \left[ \left( (2 - \alpha_1\gamma + 2\tau(\alpha_1 - 1)\gamma + \tau^2\gamma(2 - \alpha_1))\alpha_1^2 - (1 - \alpha_1)^2 g^2 \right) + \left( 2\gamma - 4\alpha_1 + 2\alpha_1^2 + 2\tau(4\alpha_1 - 2\alpha_1^2 - \gamma) \right) \times \right. \\ \left. \times (\alpha_1\gamma + (1 - \alpha_1)(1 - \gamma)g^2) + (2 - 5\alpha_1 + 4\alpha_1^2 - \alpha_1^3 + \right. \\ \left. + 2\tau(2\alpha_1 - 3\alpha_1^2 + \alpha_1^3 + \alpha_1\gamma - \gamma) + \tau^2(2\alpha^2 + \alpha^3 + 2\gamma - 2\alpha\gamma)) \times \right. \\ \left. \left. \left( \gamma^2 + (1 - \gamma)^2 g^2 \right) / \left( (\alpha_1\gamma(4 - 2\alpha_1 - \gamma)) + 1 + g^2 \sum_{i=1}^{\tau} i^2 \right) \right] \sigma_\varepsilon^2$$

**Процедура 3. Модель Брауна.**

$$\begin{aligned} \hat{a}_{0,t} &= \hat{a}_{0,t-1} + \hat{a}_{1,t-1} + (1 - \beta^3) e(t), \\ \hat{a}_{1,t} &= \hat{a}_{1,t-1} + (1 - \beta)^2 e(t), \\ e(t) &= g(t) - \tilde{g}_\tau(t), \quad \tau = 1, \end{aligned} \quad (19)$$

де параметр  $\beta$  - коефіцієнт дисконтування, що характеризує знецінення даних спостереження за одиницю часу,  $0 < \beta < 1$ ;  $e(t)$  - похибка прогнозування.

Адаптивні моделі з сезонним ефектом подаються з адитивними або мультиплікативними складовими, що накладаються на модель тренду.

Моделі першого типу описуються як :

$$g(t) = m_t + \hat{g}_t + \varepsilon_t, \quad (20)$$

де  $m_t$  - складова тренду,  $\hat{g}_t, \hat{g}_{t-1}, \dots, \hat{g}_{t-l+1}$  - адитивні коефіцієнти сезонності,  $\varepsilon_t$  - неавтокорельований шум з нульовим математичним сподіванням.

Моделі другого типу представлено у вигляді:

$$g(t) = m_t \hat{f}_t + \varepsilon_t, \quad (21)$$

де  $\hat{f}_t, \hat{f}_{t-1}, \dots, \hat{f}_{t-l+1}$  - мультиплікативні коефіцієнти сезонності;  $l$  - періодичність сезонного явища.

Як і для адаптивних моделей без сезонної складової в залежності від виду тренду та сезонної складової обчислюються поточні оцінки коефіцієнтів  $\hat{a}_{0,t}, \hat{a}_{1,t}$ .

Згідно [6] введемо поняття:

$$\hat{a}_{0,t} = \alpha_1 d_1 + (1 - \alpha_1) d_2, \quad 0 < \alpha_1 < 1 \quad (22)$$

$$\hat{a}_{1,t} = \alpha_2 (\hat{a}_{0,t} - \hat{a}_{0,t-1}) + (1 - \alpha_2) \hat{a}_{1,t-1}, \quad 0 < \alpha_2 < 1 \quad (23)$$

$$\hat{r}_t = \alpha_r \frac{\hat{a}_{0,t}}{\hat{a}_{0,t-1}} + (1 - \alpha_r) \hat{r}_{t-1}, \quad 0 < \alpha_r < 1; \quad (24)$$

$$\begin{aligned} \hat{g}_t &= \alpha_g (g(t) - \hat{a}_{0,t}) + (1 - \alpha_g) \hat{g}_{t-1}, \\ 0 &< \alpha_g < 1 \end{aligned} \quad (25)$$

$$\hat{f}_t = \alpha_f \frac{g(t)}{\hat{a}_{0,t}} + (1 - \alpha_f) \hat{f}_{t-1}, \quad 0 < \alpha_f < 1 \quad (26)$$

де  $d_1, d_2$  - коефіцієнти відповідно до адаптивної моделі;  $\hat{r}_t$  - коефіцієнт експоненціального тренду;  $\hat{g}_t$  - адитивний сезонний компонент;  $\hat{f}_t$  - мультиплікативний коефіцієнт сезонності.

Обчислення похибки прогнозування проводиться згідно:

$$\Delta = \frac{1}{\tau} \sum_{t=1}^{\tau} \frac{|g(t) - \tilde{g}_\tau(t)|}{g(t)} \cdot 100\%, \quad (27)$$

З урахуванням понять (20) – (27), нижче подано процедури отримання оцінок поточних коефіцієнтів

$\hat{a}_{0,t}$ ,  $\hat{a}_{1,t}$  та прогнозного значення  $\tilde{g}_\tau(t)$  за рядом гідрохімічних показників (1).

**Процедура 4. Адитивний лінійний тренд з адитивним сезонним ефектом.**

$$d_1 = g(t) - \hat{g}_{t-1}, \quad d_2 = \hat{a}_{0,t-1} + \hat{a}_{1,t-1} \quad (28)$$

$$\tilde{g}_\tau(t) = \hat{a}_{0,\tau} + \hat{a}_{1,\tau} + \hat{g}_{t-l+\tau} \quad (29)$$

**Процедура 5. Адитивний лінійний тренд з мультиплікативним сезонним ефектом.**

$$d_1 = g(t) / \hat{f}_{t-1}, \quad d_2 = \hat{a}_{0,t-1} + \hat{a}_{1,t-1} \quad (30)$$

$$\tilde{g}_\tau(t) = (\hat{a}_{0,\tau} + \hat{a}_{1,\tau}) \hat{f}_{t-l+\tau} \quad (31)$$

**Процедура 6. Експоненціальний тренд з адитивним сезонним ефектом.**

$$d_1 = g(t) - \hat{g}_{t-1}, \quad d_2 = \hat{a}_{0,t-1} \hat{r}_{t-1} \quad (32)$$

$$\tilde{g}_\tau(t) = \hat{a}_{0,\tau} \hat{r}_t^\tau + \hat{g}_{t-l+\tau} \quad (33)$$

**Процедура 7. Експоненціальний тренд з мультиплікативним сезонним ефектом.**

$$d_1 = g(t) / \hat{f}_{t-1}, \quad d_2 = \hat{a}_{0,t-1} \hat{r}_{t-1} \quad (34)$$

$$\tilde{g}_\tau(t) = \hat{a}_{0,\tau} \hat{f}_{t-l} + \hat{r}_t^\tau \quad (35)$$

З урахуванням процедур 4. - 7. побудовано алгоритм 3. для отримання прогнозного значення  $\tilde{g}_\tau(t)$  за рядом гідрохімічних показників (1).

#### **Алгоритм 3.**

1. Вибирається адаптивна модель прогнозування та кількість кроків прогнозування  $\tau$ .
2.  $\tau' = 1$ .
3. Якщо обрана адаптивна модель без сезонного ефекту тоді переходимо на п.4 інакше переходимо на п.8.
4. Якщо обрана модель Тейла-Вейджа, тоді задаються коефіцієнти  $\alpha_1$ ,  $\alpha_2$ , за процедурою 1 обчислюються оцінки поточних коефіцієнтів  $\hat{a}_{0,t}$ ,  $\hat{a}_{1,t}$  згідно (16).
5. Якщо обрана модель Дж.Бокса, Ч.Дженкінса, тоді задаються коефіцієнти  $\alpha_1$ ,  $\alpha_2$ ,  $\alpha_3$ , за процедурою 2 обчислюються оцінки поточних коефіцієнтів  $\hat{a}_{0,t}$ ,  $\hat{a}_{1,t}$  згідно (17) з урахуванням (18).
6. Якщо обрана модель Брауна, тоді задаються коефіцієнти  $\alpha_1$ ,  $\alpha_2$ , за процедурою 3 обчислюються оцінки поточних коефіцієнтів  $\hat{a}_{0,t}$ ,  $\hat{a}_{1,t}$  згідно (19).
7. Обчислюється прогнозне значення  $\tilde{g}_{\tau'}(t)$  згідно (15).
8. Якщо обрана модель адитивного лінійного тренду з сезонним ефектом типу (20), тоді задаються коефіцієнти  $\alpha_1$ ,  $\alpha_2$ ,  $\alpha_g$ , за процедурою 4 обчислюються оцінки поточних коефіцієнтів (22), (23) згідно (28) та прогнозне значення  $\tilde{g}_{\tau'}(t)$  знаходиться згідно (29).
9. Якщо обрана модель адитивного лінійного тренду з сезонним ефектом типу (21), тоді задаються коефіцієнти  $\alpha_1$ ,  $\alpha_2$ ,  $\alpha_f$ , за процедурою 5 обчислюються оцінки поточних коефіцієнтів (22), (23) згідно (30) та прогнозне значення  $\tilde{g}_{\tau'}(t)$  знаходиться згідно (31).
10. Якщо обрана модель експоненціального тренду з сезонним ефектом типу (20), тоді задаються коефіцієнти  $\alpha_1$ ,  $\alpha_2$ ,  $\alpha_g$ ,  $\alpha_r$ , за процедурою 6 обчислюються

оцінки поточних коефіцієнтів (22), (23) згідно (32) та прогнозне значення  $\tilde{g}_{\tau'}(t)$  знаходиться згідно (33).

11. Якщо обрана модель експоненціального тренду з сезонним ефектом типу (21), тоді задаються коефіцієнти  $\alpha_1$ ,  $\alpha_2$ ,  $\alpha_f$ ,  $\alpha_r$ , за процедурою 7 обчислюються оцінки поточних коефіцієнтів (22), (23) згідно (34) та прогнозне значення  $\tilde{g}_{\tau'}(t)$  знаходиться згідно (35).
12. Збільшується  $\tau' = \tau' + 1$ .
13. Якщо  $\tau' \leq \tau$  повторюються пункти 3-12.
14. Обчислюється похибка прогнозування у вигляді (27).
15. Закінчити роботу алгоритму.

Для вибору адекватного методу прогнозування за рядом гідрохімічних показників (1) алгоритм представлено нижче.

Для опису алгоритму введемо поняття  $\Delta_j$  - похибка прогнозування за  $j$ -ю адаптивною моделлю,  $j = \overline{1,7}$ , що обчислюється згідно (27).

#### **Алгоритм 4.**

1. Згідно алгоритму 3. проводиться прогнозування ряду (1) на  $\tau$  кроків за всіма адаптивними моделями.
2. Проводиться обчислення похибки прогнозування  $\Delta_j$  для  $j$ -ї адаптивної моделі,  $j = \overline{1,7}$ .
3. Знаходиться  $\min_j(\Delta_j)$ .
4. Модель з мінімальною похибкою прогнозування вважається адекватною для прогнозування ряду (1).
5. Закінчити роботу алгоритму.

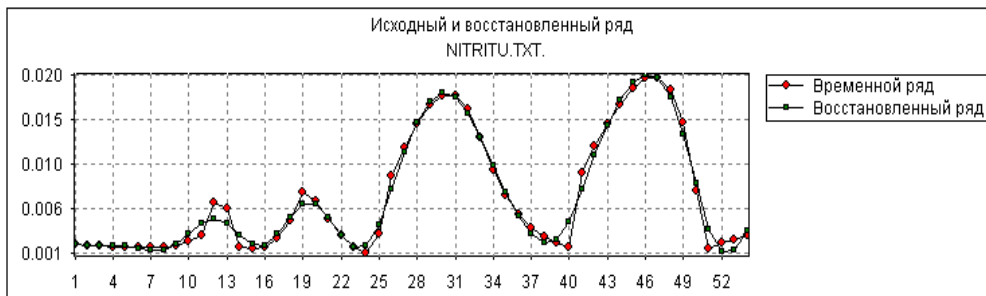
За рахунок істотного обсягу обчислень, застосування підходу передбачає автоматизацію процесу часового аналізу тому даний підхід було реалізовано за допомогою геоінформаційної системи «AquaGIS» та програмного продукту Caterpillar 2.00 [8].

Застосування представленого підходу показано частково на прикладі суміші елементів  $NO_2$ , який є одним з основних показників забрудненості води.

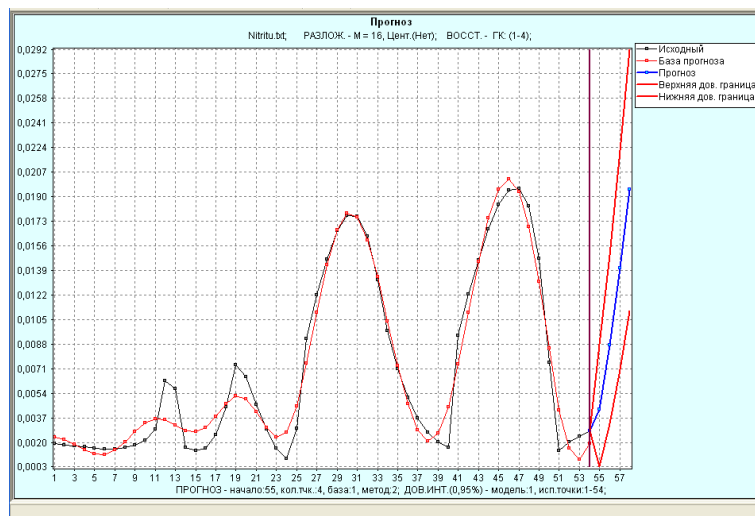
За допомогою програми Caterpillar було проаналізована структура ряду на прикладі суміші елементів  $NO_2$ , виявлено та вилучені неперіодичні коливання та побудовано прогноз на 3 кроки (квартал) (рис. 2).

Використовуючи дані про структуру ряду за  $NO_2$ , що отримано за допомогою методу SSA-«Гусениця», було проведено прогнозування ряду за адаптивними методами з сезонним ефектом (лінійний адаптивний та експоненційний тренд з мультиплікативним та адаптивним ефектом) та без сезонної складової (модель Тейла-Вейджа, Дж. Бокса, Ч. Дженкінса та Брауна) (рис. 3) з вибором адекватного методу прогнозування (табл. 1).

В результаті аналізу даних табл. 1. було зроблено висновок, що адекватного методу прогнозування у вигляді моделі лінійного аддитивного тренду з мультиплікативною сезонністю. Похибка прогнозу в середньому складає 0,007654%. За використанням даної моделі було отримано прогнози значення на квартал вперед: 11.07 - 0,0044 мг/дм<sup>3</sup>; 12.07 - 0,0061 мг/дм<sup>3</sup>; 01.08 - 0,01 мг/дм<sup>3</sup> та проведено оцінку прогнозних значень згідно ГПК ПВ. Виявлено, що на прогнозуємий квартал перевищень згідно ГПК немає ( $NO_2$  - 3.00 мг/дм<sup>3</sup>).



а)



б)

Рис.2. Проведення часового аналізу за методом SSA - «Гусениця» на прикладі  $NO_2$ : а) вияв структури ряду за даними ГХМ ПВ ; б) прогнозування за даними ГХМ ПВ.

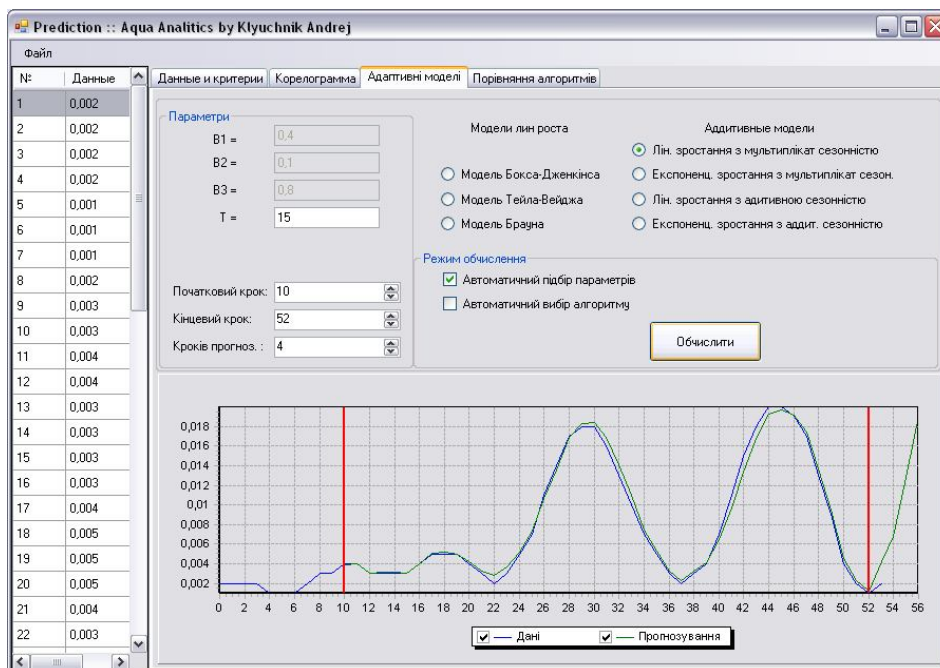


Рис. 3. Прогнозування за адитивними методами даних ГХМ ПВ на прикладі  $NO_2$ .

Таблиця 1. Вибір адекватного методу прогнозування для  $NO_2$ 

Назва методу	Параметри методу				Похибка методу $\Delta$
	T	$\alpha_1$	$\alpha_2$	$\alpha_3$	
Метод SSA – «Гусениця»	15	-	-	-	24,43465
<b>Адаптивні методи прогнозування</b>					
Модель Дж.Бокса, Ч.Дженкінса	15	0,05	0,15	0	65,53701
Модель Тейла-Вейджа	-/-	0,05	0,15	0	65,53701
Модель Брауна	-/-	0,95	0	0	76,422
Лінійний адитивний тренд з адитивною сезонною складовою	-/-	0,1	0,65	0,15	0,21685
Лінійний адитивний тренд з мультиплікативною сезонною складовою	-/-	0,4	0,1	0,8	0,007654
Експоненційний адитивний тренд з адитивною сезонною складовою	-/-	0	0	0	0,535748
Експоненційний адитивний тренд з мультиплікативною сезонною складовою	-/-	0,95	0,05	0,65	0,394016

Згідно даних, що надано у СЕС Дніпропетровського району Дніпропетровської області у 2009 році, концентрація  $NO_2$  у с. Чумаки водозабір № 1 склала на 11.07-0,0042 мг/дм<sup>3</sup>, на 12.07-0,0059 мг/дм<sup>3</sup> та на 01.08-0,008 мг/дм<sup>3</sup>.

Аналогічно було проведено часовий аналіз по всіх показниках та точках водозбору ПВ Дніпропетровського району Дніпропетровської області.

Адекватність моделі оцінювалась відповідно до  $\Delta$ . Модель з найменшою похибкою вважалась адекватною та обиралась у якості моделі для прогнозу, інші моделі також приймати участь в процесі прогнозування та працювали у режимі експерта.

Зазначена технологія в подальшому буде використана для побудови прогнозних мап стану ПВ за використанням геоінформаційної системи «AquaGIS» для оцінки динаміки змін концентрації речовин у часі.

#### ЛІТЕРАТУРА

- Шитиков В. К. Количественная гидроэкология: методы системной идентификации [Электронный ресурс] / Шитиков В.К., Розенберг Г.С., Зинченко Т.Д. – Тольятти: ИЭВБ РАН, 2003. – 463 с. – Режим доступа до книги: <http://www.ievbran.ru/kiril/Library/Book1/content0/content0.htm>.
- Вейвлет-преобразование и анализ временных рядов. – Режим доступа: <http://www.masters.donntu.edu.ua/2005/kita/piskun/library/18.htm>
- Данилов Д.Л. Главные компоненты временных рядов: метод «Гусеница» / Д.Л. Данилов, А.А. Жиглявский – 1997, СПбГУ, 308 с.
- Приставка О.П. Идентификация тренда на основе параметрических сплайнов при прогнозировании часовых рядов / О.П. Приставка, О.Г. Байбуз, Т.Г. Смельяненко // Вісник Національного університету водного господарства та природокористування. – 2006. – вип. 4(36). – ч. 2. – С. 143-151.
- Приставка О.П. Інформаційна технологія прогнозування рівня концентрації хімічних речовин / О.П. Приставка, Т.Г. Смельяненко // Актуальні проблеми автоматизації та інформаційних технологій. – Д.: Вид-во Дніпропетр. ун-ту, – 2005.– Т.9 – С.13-21.
- Приставка О.П. Інформаційна технологія обробки даних гідрогеохімічного моніторингу / О.П. Приставка, О.Г. Байбуз, Т.Г. Смельяненко // Науковий вісник НГУ, – 2006 № 3. – С. 36-41.
- Бабак В.П. Статистична обробка даних / [Бабак В.П., Білецький А.Я., Приставка О.П., Приставка П.О.]. – К.: "МІВВЦ", 2001. – 388 с.
- Лукашин Ю.П. Адаптивные методы краткосрочного прогнозирования временных рядов. – М.: Финансы и статистика, 2003. – 416 с.
- Програмная реализация метода "Гусеница". – Режим доступа: <http://www.gistatgroup.com/gus/programs.html>

пост. 19.10.09