

УДК 519.816:519.237.8

Т. Г. Ємел'яненко, Д. Л. Самарська

Дніпропетровський національний університет імені Олеся Гончара

ПОБУДОВА КОРОТКОТЕРМІНОВИХ ПРОГНОЗІВ З ВИКОРИСТАННЯМ МАТЕМАТИЧНОГО АПАРАТУ НЕЧІТКОЇ ЛОГІКИ

Запропоновано використовувати нечітку логіку з метою побудови моделей часових рядів зміни концентрації хімічних компонентів у пробах підземних вод на території гірничо-збагачувального комбінату. Розглянутий підхід покладено до основи розробленого програмного продукту *TimeSeries_FuzzyModel*, який створено у середовищі *Microsoft Visual Studio 2008* на мові програмування *C#*.

Ключові слова: *нечітка логіка, нечіткий часовий ряд, програмний продукт, прогнозування, гідрогеохімічний моніторинг.*

Предложено использовать нечеткую логику с целью построения моделей временных рядов изменения концентрации химических компонентов в пробах подземных вод на территории горно-обогатительного комбината. Рассмотренный подход положен в основу разработанного программного продукта *TimeSeries_FuzzyModel*, созданного в среде *Microsoft Visual Studio 2008* на языке программирования *C#*.

Ключевые слова: *нечеткая логика, нечеткий временной ряд, программный продукт, прогнозирование, гидрогеохимический мониторинг.*

It is proposed to use fuzzy logics for making models of hydrogeochemical time series on the territory of mountain-concentrating industrial complex. This approach was used in the developed software product *TimeSeries_FuzzyModel* which was created in *Microsoft Visual Studio 2008* on the programming language *C#*.

Key word: *fuzzy logics, fuzzy time series, software product, forecasting, hydro-geochemical monitoring.*

Вступ. Наявність у часових рядах нестаціонарності, неоднорідності, складної динаміки, відсутність інформації щодо ймовірнісних характеристик стохастичного процесу, мала кількість елементів часового ряду ускладнює та обмежує застосування класичних статистичних методів їх аналізу. Одним з можливих підходів до вирішення задачі аналізу та прогнозування в цих випадках є підхід з використанням нечіткої логіки.

Аналіз останніх досліджень і публікацій. Нечіткі часові ряди використовуються під час прогнозування даних фінансових ринків [1; 2], екологічних даних [3; 4] та в інших предметних областях. Актуальною є задача дослідження можливості застосування нечіткого підходу до прогнозування техногенного навантаження на територіях гірничо-збагачувальних комбінатів (ГЗК).

Постановка задачі. Задано часовий ряд u_t , $t = \overline{1, N}$, що містить інформацію про концентрацію хімічної речовини в пробах ґрунтових вод на техногенно навантажених територіях гірничо-збагачувальних комбінатів (ГЗК). Необхідно побудувати прогноз, тобто побудувати продовження часового ряду у вигляді

$$u_t, \quad t = \overline{1, N+l},$$

де l – довжина прогнозу.

Основний матеріал. В основі нечітких моделей лежить сукупність нечітких правил «ЯКЩО-ТО», які описують залежності між нечіткими змінними предметної області, композиційне правило виведення і спосіб обчислення значень нечітких змінних (спосіб нечіткого виведення). До системи нечіткого логічного виведення входять наступні об'єкти [5]:

- 1) сукупність нечітких продукційних правил (база правил);
- 2) набір функцій приналежності бази нечітких змінних (база змінних);
- 3) блок фаззифікації;
- 4) блок дефаззифікації;
- 5) блок виведення.

База правил містить множину логічних правил виведення, а також їх порядок (ієрархічну структуру) застосування. База нечітких змінних містить назви лінгвістичних термів та параметри їх функцій приналежності. Бази правил і нечітких змінних утворюють базу знань системи нечіткого виведення.

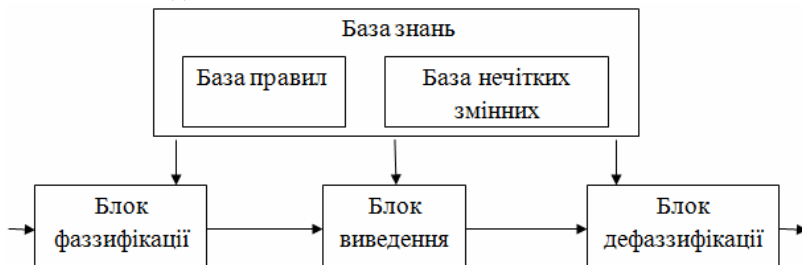


Рис 1. Система нечіткого виведення

В якості функцій приналежності використані гаусові функції, які описуються вектором параметрів $B = \{\sigma, c\}$:

$$\mu(x, B) = \exp\left(-\left(\frac{x-c}{\sigma}\right)^2\right),$$

де c – середнє значення, σ – середньоквадратичне відхилення.

Для побудови прогнозу було застосовано алгоритм нечіткого виведення Мамдані.

Алгоритм 1

1. Обчислити всі прирости для сусідніх елементів часового ряду.
2. Визначити мінімальний та максимальний прирости.
3. Поділити область визначення приростів на певну кількість інтервалів.
4. Увести в розгляд лінгвістичні змінні, які характеризують нечіткі множини.
5. Фазифікувати кожен зі знайдених приростів, тобто асоціювати його з певною нечіткою множиною в залежності від того, в який з інтервалів розбиття універсальної множини він потрапляє.
6. Розглянути попарно всі фазифіковані прирости і побудувати нечітку базу правил вигляду $A_i \rightarrow A_j$. Для кожного з приростів з метою знаходження прогнозного значення наступного приросту необхідно виконати:
 - 6.1. Знайти значення функції приналежності даного приросту X для лівих частин кожного з правил $\mu_{A_i}(X)$. Надалі розглядати ті правила нечіткої бази, для яких $\mu_{A_i}(X) > 0$.
 - 6.2. Об'єднати виходи правил логічним максимумом, тобто для правих частин кожного з правил вигляду $A_i \rightarrow A_j$ знайти $\mu'_{A_j}(X) = \max_i \mu_{A_i}(X)$.
 - 6.3. Виконати дефазифікацію, тобто знайти чітке значення наступного приросту (наприклад, як абсцису центру мас фігури, яка є агрегованою функцією приналежності).
 - 6.4. Знайти наступне значення часового ряду, додавши знайдений дефазифікований приріст до попереднього значення часового ряду.

Для підбору параметрів функцій приналежності використовується генетичний алгоритм, який описано нижче. Уведемо деякі позначення.

$$FF(S) = -\sum_{l=1}^M \left[\sum_{j=1}^m \left[\mu^{dj}(X^l, W, B, C) - \mu^{dj}(X^l) \right]^2 \right] - \text{функція відповідності};$$

K – кількість хромосом у початковому наборі; p_c – коефіцієнт схрещування.

Алгоритм 2

1. Згенерувати вхідну популяцію.
2. Знайти значення функцій відповідності $FF(S_i), i = \overline{1, K}$.
3. Визначити $\frac{K * p_c}{2}$ пар хромосом-батьків.
4. Виконати операцію схрещування кожної пари хромосом-батьків.
5. З імовірністю p_m здійснити мутацію отриманих хромосом-нащадків.
6. З отриманої популяції розміром $K + K * p_c$ хромосом відкинути $K + K * p_c$ хромосом, які мають найгірші значення $FF(S_i)$.
7. Якщо отримано хромосому, для якої $FF(S_i) = 0$, то кінець алгоритму, інакше перейти до кроку 8.
8. Якщо не вичерпано задану кількість кроків, то перейти до кроку 2; в іншому випадку хромосома, яка має найбільше значення функції відповідності $FF(S_i)$, представляє локальний оптимум. Кінець алгоритму.

Розглянуті алгоритми покладені до основи розробленого програмного продукту TimeSeries_FuzzyModel. Середовище розробки програмного продукту – Microsoft Visual Studio 2008. Мова програмування – C#. Мінімальні вимоги до системи: операційна система Windows XP/7; Framework 3.5. Програма складається з декількох модулів, кожен з яких має прив'язку до власної форми. Розроблене застосування має дружній та зручний інтерфейс для користувачів, дані наочно представлені у вигляді таблиць та графіків.

Апробація розробленого програмного продукту виконана на даних гідрогеохімічного моніторингу території Орджонікідзевського гірничо-збагачувального комбінату (ОГЗК). На рис. 2 наведено історичні дані, які є результатом вимірювання концентрації іонів магнію у воді, взятої для проб на території ОГЗК за період з грудня 1986 р. по листопад 1990 р., та побудований за описаними вище алгоритмами прогноз.

Середньоквадратична помилка дорівнює 32.66.

У табл. 1–2 наводяться обчислені значення параметрів функцій приналежності та побудована база правил нечіткого виведення.

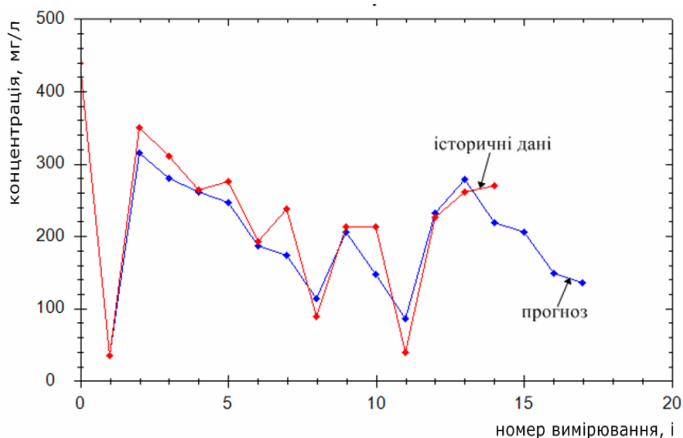


Рис. 2. Історичні дані та прогноз

Таблиця 1

Параметри функції приналежності

<i>C</i>	σ
- 404	42
- 292	96
- 195	38
- 77	52
53	52
249	60
315	60

Таблиця 2

База правил нечіткого виведення

Ліва частина правила	Права частина правила
A0	A6
A2	A4, A5
A3	A3, A4
A4	A2, A3, A4
A5	A4
A6	A3

На рис. 3 наведені графіки налаштованих функцій приналежності.

Індекси отриманих проміжків, на яких будуються функції приналежності в термінах вмісту концентрації магнію означають: 1 – дуже значне зменшення; 2 – значне зменшення; 3 – середнє зменшення; 4 – незначне зменшення; 5 – незначне збільшення; 6 – середнє збільшення; 7 – значне збільшення.

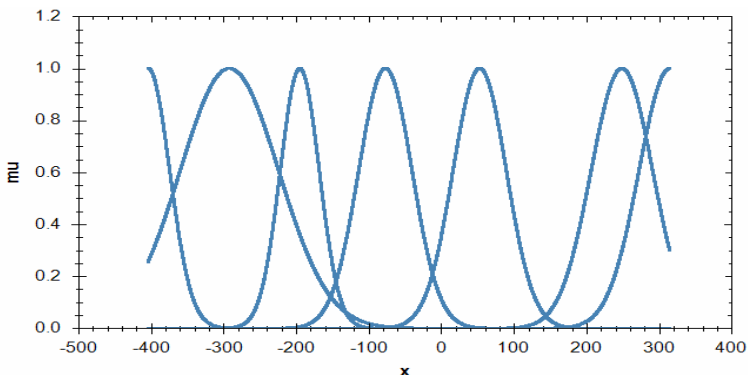


Рис. 3. Графіки функцій приналежності

Висновки. У роботі розглянута можливість застосування нечіткої логіки для аналізу та прогнозування часових рядів концентрації хімічних компонентів у пробах підземних вод на території гірничозбагачувального комбінату. Подальші дослідження можуть бути спрямовані на вдосконалення розглянутого підходу для виявлення більш складних шаблонів у нечітких часових рядах.

Бібліографічні посилання

1. Bahrepor, Majid. An adaptive ordered fuzzy time series with application to FOREX/ Majid Bahrepor, Mohammad-R. Akbarzadeh-T., Mahdi Yaghoobi, Mohammad-B. Naghibi-S // Expert systems with Applications, – № 38 (2011). – P. 475 – 485.
2. Khashei, Mehdi. A new hybrid artificial neural networks and fuzzy regression model for time series forecasting / Mehdi Khashei, Seyed Reza Hejazi, Mehdi Bijari.
3. Cheng, Ching-Hsue. Predicting daily ozone concentration maxima using fuzzy time series based on a two-stage linguistic partition method / Ching-Hsue Cheng, Sue-Fen Huang, Hia-Jong Teoh // Computers & Mathematics with Applications, – August 2011, – Vol. 62, Issue 4, P. 2016-2028.
4. Zounemat-Kermani, Mohammad. Using adaptive neuro-fuzzy inference system for hydrological time series prediction / Mohammad Zounemat-Kermani, Mohammad Teshnehlab // Applied Soft Computing, – 2008. – Vol. 8, Issue 2, P. 928 – 936.
5. **Ярушкіна, Н. Г.** Интеллектуальный анализ временных рядов: учебное пособие / Н. Г. Ярушкіна, Т. В. Афанасьева, И.Г. Перфильева. – Ульяновск; – 2010. – 320 с.

Надійшла до редколегії 23.09.11