

Міністерство освіти і науки України  
Вінницький національний технічний університет

**СНІГУР ОЛЬГА ОЛЕКСІЇВНА**

УДК 004.891.2:550.8.052

**НЕЧІТКІ МОДЕЛІ РОДОВИЩ ПІДЗЕМНИХ ВОД ТА МЕТОДИ  
ОЦІНЮВАННЯ ЇХНЬОГО СТАНУ ЗА УМОВ НЕВИЗНАЧЕНОСТІ**

01.05.02 – математичне моделювання та обчислювальні методи

**АВТОРЕФЕРАТ**

дисертації на здобуття наукового ступеня  
кандидата технічних наук

Вінниця – 2018

Дисертацією є рукопис.

Робота виконана у Вінницькому національному технічному університеті Міністерства освіти і науки України.

Науковий керівник: кандидат технічних наук, доцент  
**Кондратенко Наталія Романівна**,  
Вінницький національний технічний університет,  
професор кафедри захисту інформації

Офіційні опоненти: доктор технічних наук, професор  
**Зайченко Юрій Петрович**,  
Інститут прикладного системного аналізу  
Національного технічного університету України  
«Київський політехнічний інститут імені І. Сікорського» Міністерства освіти і науки України,  
професор кафедри математичних методів системного аналізу.

доктор технічних наук, старший науковий співробітник

**Гуляницький Леонід Федорович**,  
Інститут кібернетики імені В. М. Глушкова  
Національної академії наук України, завідувач  
відділу методів комбінаторної оптимізації та інтелектуальних інформаційних технологій.

Захист відбудеться 12 жовтня 2018 р. о 12<sup>30</sup> годині на засіданні спеціалізованої вченої ради Д 05.052.01 Вінницького національного технічного університету за адресою: 21021, м. Вінниця, вул. Хмельницьке шосе, 95, ГУК, ауд. 210.

З дисертацією можна ознайомитись у бібліотеці Вінницького національного технічного університету за адресою: 21021, м. Вінниця, вул. Хмельницьке шосе, 95.

Автореферат розіслано 7 вересня 2018 р.

Вчений секретар  
спеціалізованої вченої ради

С. М. Захарченко

## ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

**Обґрунтування вибору теми дослідження.** У багатьох сучасних дослідницьких програмах та виробничих процесах виникає необхідність побудови багатовимірних моделей складних природних об'єктів, які б із достатньою повнотою та обґрунтованістю характеризували досліджуваний об'єкт або природне явище. Такі моделі вимагають відповідного математичного апарату для роботи за умов невизначеності. Прикладами природних об'єктів такого роду є родовище підземних вод та артезіанська свердловина, яка дає змогу отримати точкові дані про характер гідрогеологічних процесів у родовищі. Для математичного опису таких систем, як правило, недостатньо сформулювати один аналітичний закон, і про характер їхньої роботи можна судити лише за вибіркою експериментальних даних, або ж на основі емпіричних знань експерта.

Родовище підземних вод як об'єкт спостережень має низку особливостей, які породжують невизначеність та унеможливають будь-які точні оцінки в процесі аналізу гідрогеологічних даних. Автоматичних чи автоматизованих систем, які б виконували функцію моделювання процесів та явищ, які виникають у свердловині, чи надавали б підтримку в прийнятті рішень, на сьогоднішній день не існує. Але підходи на основі математичного апарату нечітких множин надають потужний інструментарій для опису та моделювання залежностей у системах, що не піддаються повній формалізації. В основі нечітких моделей лежить процес формалізації причинно-наслідкових зв'язків між вхідними та вихідними змінними шляхом їхнього опису природною мовою з застосуванням нечітких множин та лінгвістичних змінних. Це досягається за допомогою функції належності, яка ставить у відповідність кожному елементу універсальної множини число з інтервалу  $[0, 1]$ , що означає ступінь належності. Поняття функції належності є узагальненням поняття характеристичної функції чіткої множини, яка оперує значеннями  $\{0, 1\}$ . Тому основні властивості та операції над нечіткими множинами являють собою узагальнення відповідних властивостей та операцій класичної теорії множин. Подальше узагальнення поняття функції належності привело до появи нечітких множин типу 2 та множин вищих порядків. Узагальнена нечітка множина визначається функціями належності, в ролі значень яких також виступають нечіткі множини. Проте побудова моделей на основі узагальнених нечітких множин пов'язана зі значною обчислювальною складністю, тому на практиці використовується їхнє інтервальне подання. Методи на основі інтервальних нечітких множин типу 2 на сьогоднішній день є поширеним засобом опису та моделювання невизначеностей, у тому числі таких, що пов'язані з недостатньою поінформованістю особи, що приймає рішення, про значення вхідних змінних моделі. Про перспективність цього підходу також свідчить кількість публікацій та різноманітних інтервальних нечітких моделей, запропонованих за останні роки.

Таким чином, існує великий клас дослідницьких та виробничих задач, що вимагають побудови математичних моделей складних природних об'єктів; до таких об'єктів належать родовища підземних вод. Щодо моделей родовищ ставиться вимога адекватного відображення об'єкта на будь-якому етапі його дослідження, навіть за умови наявності невизначеностей у вхідних даних. Це зумовлює актуальність

розробки нових моделей родовищ підземних вод та методів оцінювання їхнього стану, які дадуть змогу враховувати та моделювати невизначеності, присутні у вхідних даних, та отримувати змістовні результати на ранніх етапах дослідження родовища.

**Зв'язок роботи з науковими програмами, планами, темами.** Дисертаційну роботу виконано відповідно до напрямку наукових досліджень кафедри захисту інформації Вінницького національного технічного університету, зокрема дисертаційне дослідження пов'язане з кафедральною науково-дослідною темою № 51 К1 «Комплексна інформаційна безпека соціокомунікаційних систем», а також із науково-дослідною роботою «Розробка системи експертного оцінювання запасів підземних вод» (номер державної реєстрації 0015U007125), автор роботи є відповідальним виконавцем науково-дослідної роботи.

Основні задачі дисертації відповідають державним науково-технічним програмам, що визначені Законами України № 1977-ХІІ «Про наукову і науково-технічну діяльність», № 74/98-ВР «Про національну програму інформатизації» та державній науково-технічній програмі України за пріоритетним напрямком 6. «Інформатика, автоматизація і приладобудування».

**Мета і завдання дослідження.** *Метою дослідження* є підвищення обґрунтованості оцінок стану родовищ підземних вод шляхом розробки агрегованої моделі родовища підземних вод та методу оцінювання його стану за умов невизначеності на основі інтервальних нечітких множин типу 2.

Для досягнення мети необхідно розв'язати такі завдання.

1. Розглянути наукові першоджерела для критичного аналізу існуючих моделей складних природних об'єктів та методів оцінювання їхнього стану за умов невизначеності.

2. Розробити модель родовища підземних вод на основі інтервальних нечітких множин типу 2, яка складається з множини моделей, сформованих на основі знань експертів. Запропонувати агрегуючий критерій прийняття рішення на основі виходів множини моделей.

3. Розробити автоматизований метод виділення інформативних ознак набору експериментальних даних, що враховує знання експерта про предметну галузь.

4. Розробити вдосконалену форму методу кластеризації РСМ (Possibilistic C-Means) на основі узагальненої інтервальної оцінки за різними критеріями якості для більш повного врахування невизначеностей, пов'язаних із браком експертних знань.

*Об'єктом дослідження* є процес оцінювання стану родовищ підземних вод за умов невизначеності.

*Предметом дослідження* є нечіткі моделі та методи оцінювання стану родовищ підземних вод за умов невизначеності на основі інтервальних нечітких множин типу 2.

**Методи дослідження** базуються на основних положеннях теорії нечітких множин (для формалізації причинно-наслідкових зв'язків між вхідними та вихідними змінними моделі та побудови процесу прийняття рішень шляхом побудови нечіткого логічного висновку на основі інтервальних нечітких множин типу 2), теорії оптимізації та еволюційних обчислень (для налаштування параметрів функцій належності вхідних змінних нечітких моделей), інтелектуального аналізу даних (для

оптимізації множини вхідних змінних нечітких моделей типу 1 та 2, а також для побудови самостійної моделі прийняття рішення про зміни в характері функціонування родовища в часі).

В процесі досліджень використовувалась математична система MATLAB, програмний пакет Statistica 13, середовище розробки Microsoft Visual Studio та мова програмування C#.

**Наукова новизна отриманих результатів** полягає в тому, що:

1. Вперше запропоновано модель родовища підземних вод на основі інтервальних нечітких множин типу 2, яка, на відміну від відомих, передбачає агрегацію множини моделей, які формуються на основі знань експертів, за агрегуючим критерієм, що забезпечує підвищення обґрунтованості отриманих оцінок за умов невизначеності.

2. Вперше запропоновано метод виділення інформативних ознак набору експериментальних даних стану родовища підземних вод, який, на відміну від відомих, передбачає скорочення кількості ознак цього стану по завершенні оптимізаційної процедури, що дає змогу посилити контроль експерта за ознаками, які відкидаються, і тим самим забезпечити підвищення обґрунтованості отриманих оцінок стану родовища підземних вод.

3. Вдосконалено метод кластеризації Possibilistic C-Means шляхом введення узагальненої інтервальної оцінки за трьома критеріями якості, що забезпечує підвищення обґрунтованості оцінок перспективності родовищ підземних вод.

**Практичне значення отриманих результатів.** Практичне значення результатів, отриманих у ході дисертаційного дослідження, полягає в тому, що:

1. Створено нечітку базу знань на основі експериментальних даних артезіанських свердловин.

2. Розроблено нечітку логічну систему типу 1, яка може використовуватись не лише для оцінювання стану родовищ підземних вод, але й у низці інших задач, зокрема для оцінювання якості зображень.

3. Розроблено алгоритм отримання виходу агрегованої моделі родовища підземних вод за умов невизначеності.

4. Розроблено алгоритм автоматизованого виділення інформативних ознак набору експериментальних даних з урахуванням експертних знань.

5. Модифіковано алгоритм кластеризації Possibilistic C-Means з інтервальними ступенями належності.

6. Створено програмне забезпечення для оцінювання перспективності артезіанської свердловини за умов невизначеності.

На логічну частину розробленого програмного забезпечення отримано свідоцтва про реєстрацію авторського права на твір.

Отримані наукові результати доповнюють та розширюють відомі нечіткі моделі родовищ підземних вод і методи оцінювання їхнього стану. Розроблена математична модель та методи надають можливості для більш адекватного опису родовища на ранніх етапах роботи з ним, коли дослідник не володіє повним вектором даних, дають змогу оцінити перспективи подальшої роботи зі свердловиною та виявити випадки, коли її експлуатація пов'язана з різного роду труднощами. Запропоновані моделі артезіанських свердловин для оцінювання їхньої перспективності, поточного стану та

періодичного спостереження можуть використовуватись у довготривалих програмах вивчення родовищ підземних вод задля скорочення терміну прийняття рішень про доцільність їхньої подальшої експлуатації та заощадження необхідних для цього ресурсів.

Наукові результати дисертаційного дослідження впроваджено в практичну діяльність Правобережної геологічної експедиції ДП «Українська геологічна компанія», а також у навчальний процес кафедри захисту інформації Вінницького національного технічного університету.

**Особистий внесок здобувача.** У публікаціях, написаних у співавторстві, здобувачеві належать: метод кластеризації на основі інтервальних функцій належності типу 2 з використанням критерію якості Квона [1], запропоновано метод кластеризації на основі інтервальних функцій належності та критеріїв якості Квона, Хіе-Бені та індексу розбиття [2], виконано адаптацію методу інтервального нечіткого кластерного аналізу до прикладної задачі моніторингу стану підземних вод [3], створено нечітку базу знань на основі експериментальної вибірки зображень [8], побудовано нечіткий класифікатор для аналізу цілісності зображень [9], побудовано нечітку логічну систему на основі нечітких множин типу 1 та виконано налаштування її параметрів за допомогою генетичного алгоритму [4], виконано аналіз артезіанської свердловини як галузі прикладного застосування моделей на основі нечітких множин та запропоновано моделі в розрізі якості води, перспектив освоєння та ступеня вивченості свердловини [10], запропоновано нечітку логічну систему типу 2 для класифікації зразків води за можливим призначенням [11], запропоновано модель родовища підземних вод для оцінювання їхніх запасів на ранніх етапах розвідки [12], побудовано систему підтримки прийняття рішень на гідрогеологічних даних за умов невизначеності [15], запропоновано автоматизований метод пошуку інформативних ознак на основі генетичного алгоритму [13], запропоновано метод виділення інформативних ознак з використанням експериментальних даних та експертних знань [5], побудовано агреговану модель артезіанської свердловини для оцінювання її стану за умов невизначеності та запропоновано критерій отримання вихідного інтервалу моделі [7], [14]. У роботі [6] автору належить експериментальне підтвердження доцільності використання нечітких множин типу 2 в моделях, що оперують неповністю визначеними вхідними даними.

**Апробація матеріалів дисертації.** Основні положення і наукові результати, викладені в роботі, пройшли апробацію на науково-технічних конференціях: Міжнародній науково-технічній конференції SAIT 2013, 2014, 2015, 2016, 2017, Четвертій Міжнародній науково-практичній конференції «Методи та засоби кодування, захисту й ущільнення інформації», Міжнародній науково-практичній конференції «Інформаційні технології та комп'ютерна інженерія» та Міжнародній конференції «Контроль і управління в складних системах (КУСС-2014)». Дисертаційну роботу було представлено на науковому семінарі відділу методів комбінаторної оптимізації та інтелектуальних інформаційних технологій Інституту кібернетики імені В. М. Глушкова НАН України 12 березня 2018 р.

**Публікації.** За матеріалами роботи опубліковано 15 наукових праць, з них 7 статей в наукових фахових виданнях України, серед яких 2 у виданні, включеному до наукометричної бази Scopus та 1 – у виданні, включеному до Web of Science Core

Collection, а також 8 робіт у матеріалах і тезах доповідей міжнародних наукових конференцій. Отримано 2 свідоцтва про реєстрацію авторського права на твір.

**Структура та обсяг дисертації.** Дисертація складається зі вступу, 4 розділів, висновків, списку використаних джерел (191 найменування) і 14 додатків. Основний зміст викладено на 122 сторінках друкованого тексту, містить 33 рисунки, 9 таблиць. Загальний обсяг дисертації 210 сторінок.

## ОСНОВНИЙ ЗМІСТ РОБОТИ

У **першому розділі** розглянуто моделі родовищ підземних вод та методи оцінювання їхнього стану за умов невизначеності, що існують на сьогоднішній день. Проаналізовано існуючі класи моделей родовищ підземних вод, зокрема моделі на основі нечітких множин першого типу. Розглянуто існуючі застосування моделей на основі інтервальних нечітких множин другого типу як таких, що володіють математичним апаратом, придатним для опису та моделювання невизначеностей, присутніх у системі. Розглянуто природу невизначеностей, що можуть виникати в гідрогеологічних задачах. На основі проаналізованих відомостей визначено напрямок дослідження та сформульовано його задачі.

**Другий розділ** присвячено розробці агрегованої інтервальної нечіткої моделі родовища підземних вод, яка володіє властивістю відображати та моделювати невизначеності, в тому числі й ті, що зумовлені недостатньою поінформованістю особи, яка приймає рішення, про значення вхідних змінних.

Агрегована модель складається з множини моделей, побудованих на основі повного набору ознак та наборів ознак, виділених експертами з основного набору.

Модель має  $m$  вхідних змінних  $X^i = \langle x_1^i, \dots, x_m^i \rangle$  та одну вихідну змінну  $Y$ , причому частина значень вхідного вектора може бути відсутня внаслідок недостатньої поінформованості особи, що приймає рішення:  $x_i^z \in \emptyset, i = 1, \dots, k, k < m$ . Вхідними змінними є кількісні характеристики артезіанської свердловини, що описують особливості геологічної будови, тектонічні, кліматичні та гідрогеологічні умови, існуюче водопостачання в досліджуваному регіоні та режим експлуатації діючих водозаборів, результати обстеження експлуатаційних свердловин у районі, суміжному родовищу, а також результати дослідних робіт безпосередньо в свердловині: дані геофізичних досліджень, пробних і дослідних відкачок, режимних гідрогеологічних спостережень, та параметри, що характеризують якість підземних вод.

На основі експериментального набору значень вхідних змінних формується база знань. Правила бази знань отримуються з відомих пар  $(X_i, Y_i)$ , де вхідному вектору поставлено у відповідність лінгвістичну оцінку значення вихідної змінної  $Y$ , дану експертом. Таким чином, кожен вхідний вектор експериментальних даних генерує одне правило. Антецеденти правил утворюються заміною значення  $x_j^i$  відповідним йому нечітким термом  $A_{x_j^i}^i$ , консеквентами є терми лінгвістичної змінної  $Y$ , визначені експертом для вектора  $X^i$ :

$$R^i : \text{ЯКЩО } x_1 \in A_{x_1}^i \wedge x_2 \in A_{x_2}^i \wedge \dots \wedge x_m \in A_{x_m}^i \text{ ТО } y \in L_y^k \in \langle L_1, \dots, L_p \rangle$$

де  $x_i$  – вхідні змінні,  $A_{x_j}^i$  – терм-множини вхідних змінних, що описуються

гаусовими функціями належності вигляду  $\mu(x_i^j) = e^{-\left(\frac{x-b_i^j}{c_i^j}\right)^2}$ . Значення  $b_i^j$  прийнято рівними експериментальним значенням, значення параметру  $c_i^j$  вибрано емпірично;  $y$  – вихідна змінна,  $L_y \in \langle L_1, \dots, L_p \rangle$  – терм-множини вихідної змінної, також описуються гаусовими функціями належності.

Для покращення якості відображення моделлю навчальних даних проводиться оптимізація параметрів функцій належності. Після цього виконується перетворення результуючих функцій належності типу 1 на інтервальні функції належності типу 2 з невизначеним параметром  $c$  за таким правилом. В межах однієї ітерації для кожної з вхідних змінних значення параметру  $c$  відповідного терму функції належності змінюється в  $k$  разів. Початкове значення  $k = 1$ , на кожній наступній ітерації  $k = k + 0,001$ . Ітераційний процес повторюється доти, поки терм із максимальним покриттям результуючою функцією належності вихідної змінної залишається постійним. Таким чином отримується верхня границя інтервальної функції належності; для отримання нижньої границі виконується аналогічна процедура для  $k = k - 0,001$ .

Побудову нечіткого логічного висновку моделі забезпечує алгоритм Карніка-Менделя для інтервальних нечітких множин. Вихідна змінна приймає значення  $[y_l; y_r] \in D(Y)$ .

Інші моделі, що входять до складу агрегованої, отримуються в результаті роботи експертів із повним набором ознак з метою виділення найбільш інформативних. Кожен із експертів генерує свій набір інформативних ознак, та як наслідок окрему модель на основі інтервальних нечітких множин типу 2. База знань кожної з таких моделей отримується з загальної бази знань шляхом виключення з правил ознак, відкинутих експертом як неінформативні. У разі, коли в результаті виключення змінних із правил утворюється два або більше правила з однаковим набором антецедентів, до бази знань потрапляє правило з максимальною вагою.

Інтервальний вихід моделі за повним вхідним вектором та виходи результуючих моделей після виділення ознак експертами об'єднуються за агрегуючим критерієм  $\tilde{Y}_A = \tilde{Y}_{повн} \cap \left( \bigcup_{i=1}^n \tilde{Y}_i \right)$ , де  $\tilde{Y}_{повн}$  – інтервальний вихід моделі на основі повного набору ознак;  $\tilde{Y}_i, i = \overline{1, n}$  – вихід моделі на наборі ознак, скороченому за вказівками  $i$ -го експерта. Критерій дає інтервальну оцінку значення вихідної змінної на основі множини моделей.

Вихідна змінна носить назву «Перспективність свердловини» та показує, наскільки можливим є видобуток води в заданому місці за умови продовження гідрогеологічних досліджень.

Загальний вигляд агрегованої нечіткої моделі родовища підземних вод за умов



невизначеності показано на рис. 1.

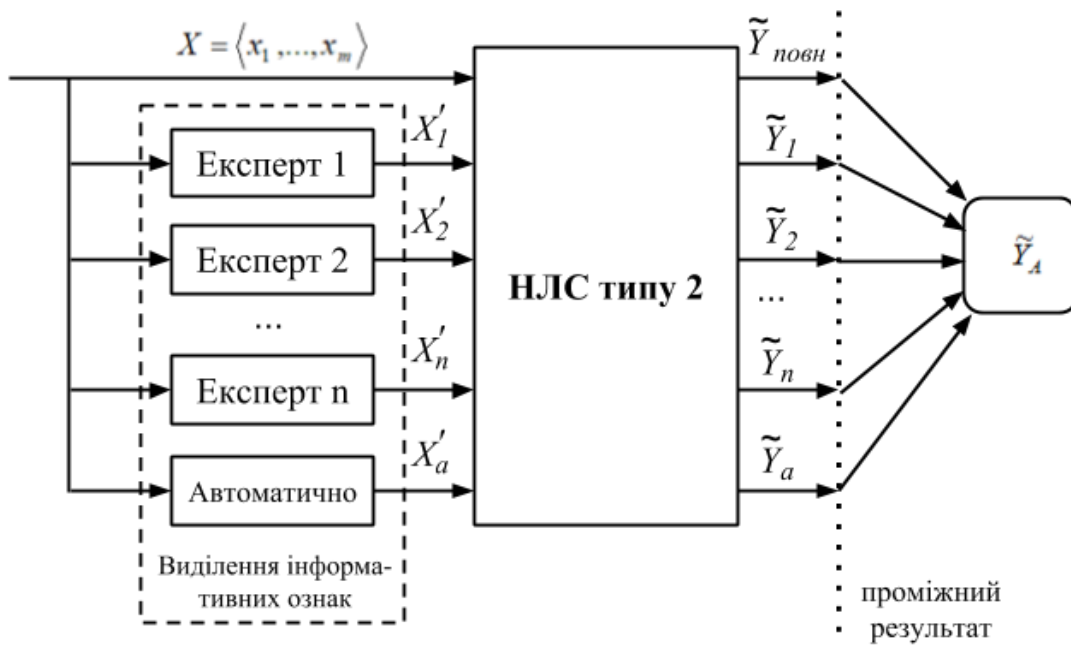


Рисунок 1 – Агрегована нечітка модель родовища підземних вод за умов невизначеності

Як джерело вхідних ознак для однієї з підмоделей запропоновано метод автоматичного вибору інформативних ознак, що виступає в ролі одного з експертів. В його основі лежить генетичний алгоритм пошуку інформативних ознак. Кожен кандидат на оптимальний розв'язок представлений хромосомою – бітовим рядком з  $k$  елементів, де  $k$  – загальна кількість усіх можливих ознак, що описують об'єкт. Якщо ознака береться до розгляду в поточному розв'язку-кандидаті, то відповідний біт хромосоми встановлюється рівним 1.

Над хромосомами кожного покоління проводяться процедури схрещування та мутації, особливості яких можуть варіюватися залежно від характеру прикладної задачі. Можливість переходу хромосоми в наступне покоління оцінюється за допомогою критерію оптимізації

$$T(C_e, C) = \sum_{i=1}^n \sum_{j=1}^c (\mu_i^j - \mu_{e_i}^j)^2 \rightarrow \min, \text{ де } \mu_{e_i} \in \{0, 1\}$$

– належність об'єкта до кластера, визначена наперед експертом для векторів навчальної вибірки,  $\mu_i \in [0, 1]$  – належність об'єкта до кластера за виходом синтезованої моделі. Для обчислення критерію  $T$  використовується розбиття  $C_i$ , побудоване на основі поточної хромосоми, та  $C_{e_i}$ , задане експертом для набору даних

$$X \text{ на основі Евклідової відстані в просторі зразків } d_E(x_1, x_2) = \sqrt{\sum_{i=1}^k (x_1 - x_2)^2}.$$

По завершенні оптимізаційної процедури проводиться кластеризація вибраних ознак методом  $c$ -means у просторі, осями якого виступають зразки  $x_1, \dots, x_n$ . Елементами кластерів у цьому випадку виступають ознаки. Відстань між двома векторами ознак  $p_1$ , та  $p_2$  також визначається за Евклідовою метрикою:

$d_E(p_1, p_2) = \sqrt{\sum_{j=1}^n (x_{j1} - x_{j2})^2}$ . За отриманим розбиттям обираються ознаки, найближчі

до центрів кластерів:  $\mu_{ij} \geq \eta$ , – таким чином усуваючи ознаки, що впливають на прийняте рішення однаковим чином.

Запропонований метод дає змогу виявити зв'язки між вхідними та вихідним параметрами моделі, і за на основі цих зв'язків скоротити простір ознак за рахунок таких, що не чинять впливу на значення вихідного параметру, або впливом яких можна знехтувати. Метод інтегрує емпіричні спостереження (навчальний набір даних) та експертні знання (еталонне розбиття, значення порогів, контроль експерта за відкиданням ознак, визначених як надлишкові.)

Розроблена агрегована модель допускає існування альтернативних підмоделей на основі інших методів прийняття рішень або формальних процедур, що вже існують для розв'язання поставленої задачі. В задачі оцінювання стану артезіанської свердловини однією з таких моделей є процедура, якою послуговуються дослідники-гідрогеологи на сьогоднішній день – перевірка відповідності кількісних параметрів складу води нормам, встановленим законодавством.

Структура та математичний апарат, що лягає в основу альтернативних моделей, диктується прикладним застосуванням. У випадку її використання в задачі оцінювання стану артезіанської свердловини альтернативна модель слугує для моніторингу змін стану свердловини з плином часу шляхом контролю зміни положення об'єкта-свердловини відносно компактних груп свердловин навчальної вибірки. Модель застосовується на завершальному етапі проведення геологорозвідувальних робіт та в подальшому процесі моніторингу, і має за основу процедуру прийняття рішень на основі методу кластерного аналізу.

У **третьому розділі** запропоновано альтернативну модель родовища підземних вод в описуваній системі моделей, побудовану на основі модифікованого методу кластеризації РСМ з інтервальним виходом. Удосконалення методу РСМ полягає в розширенні його можливостей шляхом введення інтервальних ступенів належності, що дозволяє враховувати та моделювати невизначеності, пов'язані з браком експертних знань. Математичний опис цих невизначеностей стає можливим завдяки поданню рівня нечіткості  $m$  в інтервальній формі шляхом виявлення його граничних значень за рядом критеріїв якості кластеризації. Для цього використано критерії якості кластеризації Квона, Хіе-Бені та індекс розбиття. Аналізуючи значення цих критеріїв для різних значень  $m$ , будуються інтервали зміни значення рівня нечіткості за кожним критерієм:  $\tilde{m}_K = [m_{Kl}; m_{Kr}]$ ,  $\tilde{m}_{XB} = [m_{XB1}; m_{XB2}]$  та  $\tilde{m}_{SC} = [m_{SC1}; m_{SC2}]$ , після чого вони об'єднуються за правилом:

$$\tilde{m} = \tilde{m}_K \cup \tilde{m}_{XB} \cup \tilde{m}_{SC} = [\min(m_{Kl}, m_{XB1}, m_{SC1}); \max(m_{Kr}, m_{XB2}, m_{SC2})]$$

Інтервали зміни рівня нечіткості, отримані в результаті цієї операції, можна використовувати напряму для розрахунку ступенів належності.

Удосконалена форма методу дає узагальнену оцінку за трьома критеріями якості, що дозволяє уникнути помилкових результатів, пов'язаних із внутрішніми

особливостями того чи іншого єдиного критерію. Інтервальні ступені належності дають змогу враховувати та моделювати невизначеності, пов'язані з браком експертних знань, характерним для моделей на основі навчання без учителя, таких як кластерний аналіз. Врахування та моделювання невизначеностей, закладених у вихідному наборі даних, дає змогу помітити тенденції та зміни в характері процесів, що протікають у родовищі, на стадії їхнього формування. Крім того, перевагою цього підходу є можливість працювати з кількома вхідними векторами одночасно. Галузями потенційного застосування моделі є програми поточного спостереження та підтримки вже сформованих систем, таких як поточний контроль показників набору вже існуючих артезіанської свердловини.

Запропоновано застосування моделі кластеризації на основі інтервальних ступенів належності для візуалізації даних спостережень за родовищем підземних вод у часі з метою раннього виявлення негативних тенденцій та закономірностей. Модель зображено на рис. 2, а процес моделювання відбувається в такій послідовності.

1. Визначити параметри розбиття навчального набору  $(X_n, Y_n)$  на кластери відповідно до методу, описаного вище, а саме значення критеріїв якості кластеризації  $K(c, m)$ ,  $XB(c, m)$  та  $SC(c, m)$ , а також положення центрів кластерів  $\langle c_1, \dots, c_p \rangle$  та значення рівня нечіткості  $\tilde{m}$ . Обчислювати остаточні значення ступенів належності зразків навчального набору до отриманих кластерів немає необхідності.
2. На останньому кроці методу інтервальної нечіткої кластеризації обчислити значення ступенів належності  $\langle \tilde{\mu}_1, \dots, \tilde{\mu}_p \rangle$  до кластерів для зразка, що являє собою вектор  $W^z$  параметрів контрольованої свердловини:

$$\mu_{ij} = \frac{1}{1 + \left( \frac{d_{ij}^2}{\eta_j} \right)^{\frac{1}{m-1}}}$$

В загальному випадку векторів може бути більш ніж один; в такому випадку йдеться про «підміну» навчального набору даних тестовим та обчислюються значення ступенів належності  $\langle \tilde{\mu}_{11}, \dots, \tilde{\mu}_{np} \rangle$  для всіх векторів тестового набору  $(X_m, Y_m)$ . Оскільки для оцінювання позиції кожного з контрольованих зразків достатньо лише обчислити ступені належності, на обсяг тестового набору даних, що може оброблятися, не накладається обмежень; в загальному випадку він може перевищувати обсяг навчального набору.

3. Обчислити нові значення критеріїв якості кластеризації  $K'(c, m)$ ,  $XB'(c, m)$  та  $SC'(c, m)$  на об'єднанні навчального набору даних та досліджуваної свердловини, або на тестовому наборі даних. Це дає змогу оцінити, чи спотворюють дані тестового набору еталонне розбиття, отримане для навчального набору, та визначити кількісну міру цих спотворень.

4. Остаточне рішення приймається за ступенями належності точки, що характеризується вектором  $W^z$ , до кожного з  $s$  утворених кластерів, а саме за верхньою границею інтервалу. У разі, коли верхня границя інтервальних ступенів належності однакова для двох або більше кластерів, свердловину слід віднести до кластера з меншою шириною інтервалу ступеня належності.

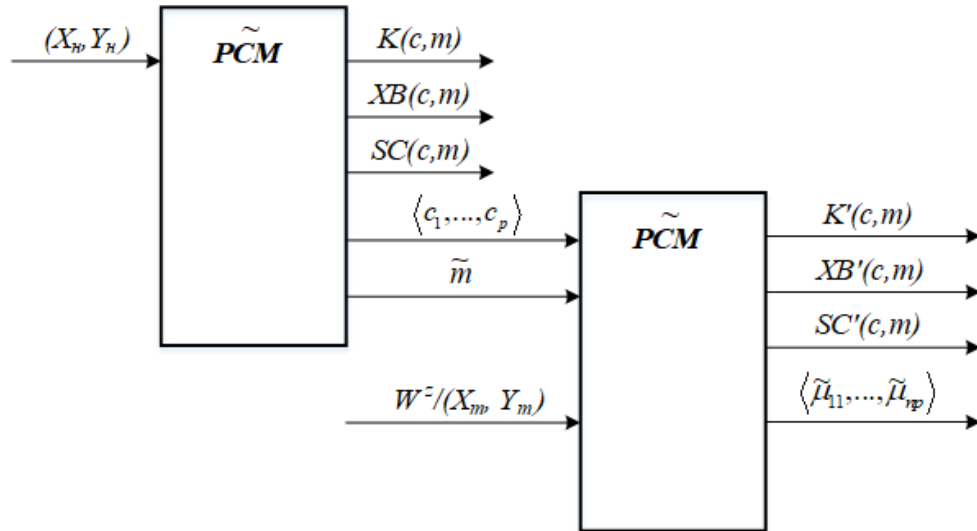


Рисунок 2 – Модель на основі інтервального нечіткого кластерного аналізу для оцінювання перспективності родовища підземних вод

Ширина отриманих інтервалів ступенів належності може давати додаткову інформацію про невизначеність, пов'язану з прийнятим рішенням. Динаміка зміни інтервальних значень ступенів належності може сигналізувати про дію негативних факторів на етапі їхнього формування та вказати на необхідність проведення додаткових досліджень.

**Четвертий розділ** присвячено дослідженню функціональних можливостей запропонованих моделей та методів у задачі оцінювання стану артезіанської свердловини на різних етапах дослідження, а також порівнянню отриманих результатів з отриманими на основі загальновідомих підходів.

Моделі, запропоновані в попередніх розділах, реалізовано в задачі оцінювання перспективності артезіанської свердловини. Проведено обробку архівних даних Правобережної геологічної експедиції та сформовано навчальну і тестову вибірки даних. Досліджено роботу моделі на основі нечітких множин типу 1 в поставленій задачі. Сформовано базу правил на основі експериментальних даних попередніх досліджень артезіанських свердловин (табл. 1).

Для покращення якості відображення моделлю експериментальних даних проведено її навчання за допомогою генетичного алгоритму. Значення параметрів та графічне подання функцій належності наведено в табл. 2.

Таблиця 1 – Фрагмент бази правил

	R <sup>1</sup>	R <sup>2</sup>	R <sup>3</sup>	R <sup>4</sup>	R <sup>5</sup>	R <sup>6</sup>	R <sup>7</sup>	R <sup>8</sup>	R <sup>9</sup>	R <sup>10</sup>	R <sup>11</sup>	R <sup>12</sup>	R <sup>13</sup>	R <sup>14</sup>	R <sup>15</sup>	R <sup>16</sup>	R <sup>17</sup>	R <sup>18</sup>	R <sup>19</sup>	R <sup>20</sup>
x <sub>1</sub>	С	В	Н	С	С	С	С	С	С	С	С	Н	В	В	В	С	С	С	С	Н
x <sub>7</sub>	В	В	В	В	С	С	С	В	В	В	С	В	В	В	В	В	В	В	В	В
x <sub>11</sub>	Н	Н	Н	Н	Н	Н	Н	Н	Н	Н	В	Н	В	В	В	В	Н	В	Н	В
...																				
y	В	В	В	В	В	В	Д	Д	Д	Д	Д	Д	Д	Д	Н	Н	Н	Н	В	Н

Таблиця 2 – Параметри функцій належності типу 1

Змінна	Назва змінної	Терми					
		Н		С		В	
		<i>b</i>	<i>c</i>	<i>b</i>	<i>c</i>	<i>b</i>	<i>c</i>
x <sub>1</sub>	Віддаленість від населених пунктів, км	0	15,58	24,7	15,62	50	15,69
x <sub>7</sub>	Концентрація радону (середня по існуючих свердловинах на даній території), Бк/дм <sup>3</sup>	0	23	49,5	22,84	101,1	22,73
...							
y	Перспективність свердловини	0	1,5	5,65	1,5	10	1,5

Виявлено, що апроксимуючі властивості моделі на основі нечітких множин типу 1 суттєво погіршуються за умови недостатньої поінформованості особи, що приймає рішення, про значення вхідних змінних (табл. 3). Так, у наведеному фрагменті в результаті повного гідрогеологічного дослідження першу свердловину охарактеризовано експертом як недостатньо перспективну, а свердловини 8-10 – як такі, що мають добрі перспективи для видобутку питної води. На першому ж етапі, коли відомі значення лише 7 ознак із 24, для свердловин 8-10 вихід моделі відповідає терму «Достатня». Помилки на третьому етапі, коли відомі значення всіх 24 ознак

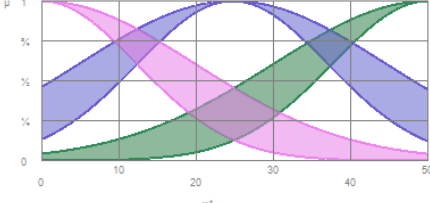
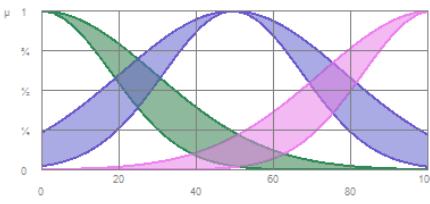
(свердловини 1 та 10), можна пояснити недосконалістю навчальної вибірки, зокрема її недостатнім обсягом. В задачах, пов'язаних із гідрогеологічними дослідженнями, цей фактор є додатковим джерелом невизначеності, оскільки отримання даних пов'язане зі значними витратами матеріальних та людських ресурсів.

Таблиця 3 – Результати роботи моделі типу 1 за умови недостатньої поінформованості особи, що приймає рішення, про значення вхідних змінних

№	Вихід моделі			Оцінка експерта
	1-й етап (7 ознак)	2-й етап (14 ознак)	3-й етап (24 ознаки)	
1	1,38 / Недостатня	4,27 / Достатня	3,94 / Достатня	1 / Недостатня
...				
8	7,54 / Достатня	8,28 / Висока	8,39 / Висока	8 / Висока
9	6,74 / Достатня	8,36 / Висока	8,2 / Висока	9 / Висока
10	5,94 / Достатня	6,6 / Достатня	7,77 / Достатня	9 / Висока

В табл. 4 наведено значення параметрів та графічне подання функцій належності термів перших двох вхідних змінних після перетворення їх в інтервальну форму.

Таблиця 4 – Інтервальні функції належності

Змінна	Назва змінної	Терми					
		Н		С		В	
		$b$	$c$	$b$	$c$	$b$	$c$
$x_1$	Віддаленість від населених пунктів, км	0	[12,28; 19,84]	24,7	[12,31; 19,88]	50	[12,37; 19,98]
							
$x_7$	Концентрація радону (середня по існуючих свердловинах на даній території), Бк/дм <sup>3</sup>	0	[18,12; 29,27]	49,5	[18; 29,07]	101,1	[17,91; 28,93]
							
...							

Перевірено роботу моделі на основі інтервальних нечітких множин типу 2 для оцінювання перспективності артезіанської свердловини. Показано роботу моделі на ранніх етапах дослідження, коли особа, що приймає рішення, не володіє інформацією

про значення багатьох вхідних змінних.

Показано реалізацію агрегованої нечіткої моделі родовища підземних вод, підмоделі якої утворені наборами інформативних ознак, запропонованих експертами. Набори містять 10 та 11 вхідних змінних відповідно. Третій набір із 12 змінних отримано за допомогою автоматизованого методу виділення інформативних ознак. Інтервальні функції належності, отримані на цьому наборі, істотно відрізняються від функцій, отриманих на повному наборі вхідних змінних (табл. 5).

Таблиця 5 – Інтервальні функції належності на скороченому наборі змінних

Змінна	Назва змінної	Терми					
		Н		С		В	
		$b$	$c$	$b$	$c$	$b$	$c$
x <sub>7</sub>	Концентрація радону (середня по існуючих свердловинах на даній території), Бк/дм <sup>3</sup>	0	[21,11; 24,01]	49,5	[20,96; 23,84]	101,1	[20,86; 23,72]
		...					
x <sub>14</sub>	Вміст нітратів (середній у прилеглих свердловинах), мг/л	0	[0,41; 0,46]	1,47	[0,41; 0,46]	2,99	[0,42; 0,48]
		...					

Агрегована модель показує змістовні результати як на завершальному, так і на проміжних етапах ГРР. Результати, показані в таблиці 6, дають змогу зробити висновок про доцільність подальших робіт на ділянці.

Не зважаючи на порівняно значну ширину інтервалу значень вихідного параметра, вже на початку геологорозвідувальних робіт можна зрозуміти, що подальша робота на вибраній ділянці має досить широкі перспективи. Із плином часу в процесі гідрогеологічного дослідження з'являються дані за рештою показників, і висновок уточнюється, а ширина інтервалу невизначеності зменшується. У разі, коли на одному з початкових етапів ГРР отримано вихід моделі, що тяжіє до недостатнього значення перспективності, можуть бути проведені додаткові дослідження та прийняте рішення про припинення робіт, що дозволить зберегти суттєві матеріальні та людські ресурси.

Таблиця 6 - Вихід моделі на трьох етапах гідрогеологічного дослідження

№ етапу ГРР	Експерт 1		Експерт 2		Автоматизований метод		Повний вектор				Агрегований вихід
	Визн. ознак/всього	Вихід моделі	Визн. ознак/всього	Вихід моделі	Визн. ознак/всього	Вихід моделі	Визн. ознак/всього	Вихід моделі	Вихід альтерн. моделі 1	Вихід альтерн. моделі 2	
1	4/10	[4,3; 5,98]	5/11	[6,35; 6,51]	4/12	[5,09; 5,92]	7/24	[7,31; 7,62]	x	x	[7,32; 7,7] Д, В
2	6/10	[3,4; 4,45]	7/11	[4,39; 4,57]	6/12	[6,62; 6,69]	14/24	[5,37; 6,15]	x	x	[5,37; 6,15] Д
3	10/10	[5,11; 5,42]	11/11	[4,78; 5,22]	12/12	[5,83; 6,02]	24/24	[3,9; 5,71]	Д:[0; 0,565]	Достатня	[4,78; 5,71] Д

Практичну користь розроблених моделей та методів із точки зору підвищення обґрунтованості отриманих оцінок стану родовища досліджено за допомогою

критеріїв  $RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - y_i^e)^2}$  та  $MaxAE = \max(|y_i - y_i^e|)$ , а також показника

$Z(k) = \frac{\sum_{i=1}^N k_i}{10N}$ , де  $k_i$  – ширина інтервалу значення вихідної змінної,  $N$  – кількість

векторів набору даних.  $Z(k)$  – це характеристика ширини інтервалу вихідного значення як оцінки рівня невизначеності, пов'язаної з отриманим виходом моделі. Ця інформація є важливою для винесення обґрунтованої оцінки стану родовища особою, що приймає рішення. Отримані оцінки наведено в табл. 7.

Таблиця 7 – Оцінки якості нечітких моделей родовища підземних вод на тестовій вибірці

Модель	1-й етап			2-й етап			3-й етап		
	<i>RMSE</i>	<i>MaxAE</i>	<i>Z(k)</i>	<i>RMSE</i>	<i>MaxAE</i>	<i>Z(k)</i>	<i>RMSE</i>	<i>MaxAE</i>	<i>Z(k)</i>
Нечітка модель типу 1	1,17	3,06	-	1,42	3,63	-	1,35	3,85	-
Нечітка модель типу 2, повний вектор	0,68	1,69	0,104	0,74	1,93	0,168	0,84	1,96	0,239
Нечітка модель типу 2, набір ознак експерта 1	1,42	5,29	0,035	1,36	5,02	0,04	1,31	4,86	0,033
Нечітка модель типу 2, набір ознак експерта 2	1,59	5,81	0,02	1,44	4,88	0,011	1,42	4,81	0,014
Нечітка модель типу 2, набір ознак автоматизованого методу	1,43	5,37	0,018	1,09	4,16	0,015	1,21	2,54	0,018
Агрегована нечітка модель	0,98	1,56	0,085	1,01	1,94	0,102	1,26	2,22	0,09



Таким чином, агрегована нечітка модель родовища підземних вод на ранніх етапах ГРР більш якісно відображає досліджуваний об'єкт у порівнянні з нечіткою моделлю типу 1. Так, на першому етапі гідрогеологічного дослідження значення критеріїв якості складають  $RMSE = 0,98$ ;  $MaxAE = 1,56$ , в той час як для нечіткої моделі типу 1 за тих самих умов  $RMSE = 1,17$ ;  $MaxAE = 3,06$ . Агрегована модель забезпечує значення  $Z(k)$  на рівні  $0,011 - 0,239$  залежно від ступеня поінформованості особи, що приймає рішення, про значення вхідних змінних моделі. В цілому можна стверджувати, що значення  $Z(k) \leq 0,4$ , як правило, забезпечує кількість інформації, достатню для кваліфікованого експертного оцінювання свердловини особою, що приймає рішення. Одержані результати свідчать про можливість отримання прийнятної оцінки стану свердловини на 2-му етапі ГРР, а в деяких випадках і на 1-му, тобто ще до початку буріння.

Часові затрати на навчання моделей на основі нечітких множин типу 1 складають 3,88 год (в середньому за 10 запусків), на побудову нечіткого логічного висновку – в середньому 513 мс на кожен вектор. Перетворення функцій належності в інтервальну форму виконується в середньому 21,22 с для випадку 24 вхідних змінних. Отримання інтервального виходу моделі на основі нечітких множин типу 2 (агрегована модель) триває в середньому 0,93 мс, тобто отримання результатів можливе в режимі реального часу. Що стосується порівняно значних часових затрат на побудову моделі, то зважаючи на тривалість гідрогеологічних досліджень, яка може складати від кількох місяців до кількох років, цей показник також знаходиться в допустимих межах. Час виконання всіх обчислень наведено для ПК Lenovo 62718QG (Intel Core i5-3230M, 2.60 GHz, 8 GB RAM).

Окрім інтервального виходу НЛС, результати також містять виходи двох альтернативних моделей. Альтернативна модель 1 – модель прийняття рішення на основі інтервального нечіткого кластерного аналізу. Альтернативна модель 2 – рішення, прийняте виходячи з гранично допустимих значень параметрів, контрольованих згідно з законодавством України.

У розгорнутому вигляді вихід альтернативної моделі на основі методу інтервальної нечіткої кластеризації виглядає як показано в табл. 8 та дає змогу оцінити не лише ступені належності кожного зразка до відповідного кластера, але й міру невизначеності, пов'язану з цією оцінкою.

Періодичний контроль цих величин дозволяє вчасно виявити зсув відповідного зразка в бік іншого кластера з характеристиками, відмінними від тих, якими описувався зразок на момент початку ведення спостережень. Так, у наведеному прикладі тестова вибірка імітує дані частини свердловин, що складають навчальну вибірку, в деякий момент часу, скажімо, за 10 років після введення їх в експлуатацію. Свердловина 4 тестової вибірки імітує поступове підвищення радіоактивності ( $x_{65}=219,5$  порівняно з  $x_{65}=53,92$  для тієї самої свердловини в навчальній вибірці). Можна помітити, що на тестовій вибірці для цієї свердловини модель більше не вказує однозначно на один із кластерів, як це відбувається на навчальній вибірці. Для порівняння, свердловина 4 навчальної вибірки належить до кластера «Висока» зі ступенем належності  $[0,689; 1]$ , до кластера «Достатня» -  $[0,542; 1]$ , до кластера «Недостатня» -  $[0; 0,418]$ . Тепер же ступені належності до всіх трьох кластерів

невисокі ( $<0,5$ ), характеризуються значною шириною інтервалу ( $0,404 - 0,454$ ), і тяжіють до кластера «Недостатня».

Таблиця 8 – Вихід альтернативної моделі на основі інтервальної нечіткої кластеризації

№	Кластер 1 (Висока)		Кластер 2 (Достатня)		Кластер 3 (Недостатня)		Результат / ширина інтервалу	Оцінка експерта
	Ліва границя	Права границя	Ліва границя	Права границя	Ліва границя	Права границя		
1	0,532	1	0,395	0,48	0	0,468	Висока / 0,468	Висока
2	0,595	1	0,526	1	0	0,488	Висока / 0,405	Висока
3	0,615	1	0,552	1	0	0,51	Висока / 0,385	Висока
4	0	0,404	0	0,415	0	0,454	Недостатня / 0,454	Недостатня
...								

Таким чином, хоча вихід моделі не завжди дає однозначну відповідь щодо належності свердловини до одного з трьох класів, він є достатньо змістовним для того, щоб звернути увагу дослідника на процеси, що відбуваються в цій свердловині, та вжити заходів для попередження впливу негативних факторів.

Що стосується часу виконання, то найдовшим етапом є визначення границь інтервалу ступеня невизначеності. Аналіз розбиттів зі значеннями  $m$  від 1 до 5 з кроком 0,01 виконується в середньому за 20,25 с. Маючи інтервал зміни значень рівня нечіткості, інтервальні значення ступені належності знаходяться в середньому за 6,01 с для 20 вхідних векторів та 2 с для 10 вхідних векторів.

На основі запропонованих моделей та методів створено програмне забезпечення для оцінювання перспективності артезіанської свердловини.

## ВИСНОВКИ

Роботу присвячено підвищенню обґрунтованості оцінок стану родовищ підземних вод шляхом розробки агрегованої моделі родовища підземних вод та методу оцінювання його стану за умов невизначеності на основі інтервальних нечітких множин типу 2. У ході дослідження отримано такі результати.

1. Досліджено існуючі моделі родовищ підземних вод та методи оцінювання їхнього стану за умов невизначеності. Виділено такі характерні риси задач, пов'язаних із моделюванням родовищ підземних вод, як багатовимірність, велика роль експертних оцінок у процесі прийняття рішення та суттєві затрати часу, необхідні для отримання значень багатьох вхідних змінних. З огляду на ці риси, особливу увагу приділено моделям та методам на основі нечітких множин.

2. Розроблено агреговану нечітку модель родовища підземних вод на основі інтервальних нечітких множин типу 2, яка складається з множини моделей, сформованих на основі знань експертів. Виявлено властивість моделі отримувати змістовний результат за умов недостатньої інформованості особи, що приймає рішення, про значення вхідних змінних, що дає змогу застосовувати її на будь-якому

етапі дослідження родовища. Показано, що отримана модель забезпечує більшу обґрунтованість прийнятих рішень про стан родовища. Це досягається, з одного боку, наявністю додаткової інформації у вигляді ширини інтервалу вихідного значення. Ця інформація має важливе значення для винесення обґрунтованої оцінки стану родовища особою, що приймає рішення. Отримана модель забезпечує значення  $Z(k)$  на рівні  $0,011 - 0,239$  залежно від ступеня поінформованості особи, що приймає рішення, про значення вхідних змінних моделі. З іншого боку, агрегована нечітка модель родовища підземних вод на ранніх етапах ГРР більш якісно відображає досліджуваний об'єкт у порівнянні з нечіткою моделлю типу 1. Так, на першому етапі гідрогеологічного дослідження значення критеріїв якості складають  $RMSE = 0,98$ ;  $MaxAE = 1,56$ , в той час як для нечіткої моделі типу 1 за тих самих умов  $RMSE = 1,17$ ;  $MaxAE = 3,06$ .

3. Розроблено автоматизований метод виділення інформативних ознак набору експериментальних даних, що враховує знання експерта про предметну галузь. Показано, яким чином використання методу в межах агрегованої нечіткої моделі родовища дає змогу надати експерту більші можливості з контролю над набором ознак, що виключаються з розгляду, та запобігти виключенню важливих ознак. Це також сприяє підвищенню обґрунтованості результуючої оцінки стану родовища підземних вод.

4. Розроблено вдосконалену форму методу кластеризації РСМ на основі узагальненої інтервальної оцінки за різними критеріями якості для більш повного врахування невизначеностей, пов'язаних із браком експертних знань. Виявлено, що введення інтервальних ступенів належності дає змогу враховувати та моделювати невизначеності, пов'язані з браком знань про досліджуваний об'єкт. Узагальнена оцінка за трьома критеріями якості кластеризації дає змогу уникнути помилкових результатів, пов'язаних із внутрішніми особливостями того чи іншого єдиного критерію. Показано, як застосування розробленого методу дає змогу виявляти негативні тенденції на ранніх стадіях їхнього формування.

Для практичного використання отриманих наукових результатів у виробничій задачі оцінювання перспективності артезіанської свердловини створено спеціальне програмне забезпечення.

Наукові результати дисертаційного дослідження впроваджено в практичну діяльність Правобережної геологічної експедиції ДП «Українська геологічна компанія» та в навчальний процес кафедри захисту інформації Вінницького національного технічного університету.

## СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

- [1] Н. Р. Кондратенко, та О. О. Снігур, «Використання інтервальних функцій належності в задачах кластеризації даних соціального характеру», *Системні дослідження та інформаційні технології*, № 3. с. 20-29, 2012.
- [2] Н. Р. Кондратенко, та О. О. Снігур, «Інтервальна нечітка кластеризація на основі альтернативних критеріїв якості», *Наукові вісті НТУУ «КПІ»*, № 2. с. 59-66, 2012.

- [3] Н. Р. Кондратенко, та О. О. Снігур, «Інтервальний нечіткий кластерний аналіз для моніторингу стану артезіанської свердловини», *Радіоелектроніка, інформатика, управління*, №4, с. 77-84, 2017.
- [4] Н. Р. Кондратенко, та О. О. Снігур, «Нечітке моделювання в задачі оцінки якості зображень лазерних плям», *Системні дослідження та інформаційні технології*, №3, с. 7-16, 2014.
- [5] Н. Р. Кондратенко, та О. О. Снігур, «Еволюційний пошук інформативних ознак із залученням експерта в задачі оцінки якості артезіанської води», *Вісник Вінницького політехнічного інституту*, № 3, с. 96-101, 2015.
- [6] N. Kondratenko, and O. Snihur, «A Method of Building Type-2 Fuzzy Logic Systems in Multidimensional Objects Identification Problems», *Eastern-European Journal of Enterprise Technologies*, V. 3/4 (87), p. 38-45, 2017.
- [7] N. Kondratenko, and O. Snihur, «Interval Fuzzy Modeling of Complex Systems under Conditions of Input Data Uncertainty», *Eastern-European Journal of Enterprise Technologies*, V. 4/4 (82), p. 20-28, 2016.
- [8] Н. Р. Кондратенко, В. А. Лужецький, та О. О. Снігур, «Аналіз зображень із використанням нечітких хеш-функцій», на *IV Міжнар. наук.-практ. конф. Методи та засоби кодування, захисту й ущільнення інформації*, Вінниця, 2013, с. 308-309.
- [9] Н. Р. Кондратенко, та О. О. Снігур, «Нечітке моделювання в задачах аналізу цілісності зображень», на *15-й Міжнар. наук.-техн. конф. SAIT 2013*, Київ, 2013, с. 287.
- [10] Н. Р. Кондратенко, та О. О. Снігур, «Нечітка модель оцінки запасів підземних вод» на *міжнар. наук.-практ. конф. Інформаційні технології та комп'ютерна інженерія*, Вінниця, 2014, с. 116–117.
- [11] Н. Р. Кондратенко, та О. О. Снігур, «Нечітке моделювання в задачах гідрогеології», на *16-й Міжнар. наук.-техн. конф. SAIT 2014*, Київ, 2014. с. 224-225.
- [12] Н. Р. Кондратенко, та О. О. Снігур, «Нечітка логічна система типу 2 для оцінювання перспективності артезіанської свердловини», на *17-й Міжнар. наук.-техн. конф. SAIT 2015*, Київ, 2015, с. 157.
- [13] Н. Р. Кондратенко, та О. О. Снігур, «Еволюційний пошук із кластеризацією для виділення інформативних ознак у задачі оцінки якості питної води», на *XII міжнар. конф. Контроль і управління в складних системах (КУСС-2014)*, Вінниця, 2014, с. 208.
- [14] Н. Р. Кондратенко, та О. О. Снігур, «Моделювання багатовимірних залежностей на основі нечітких множин типу 2 на недовизначених гідрогеологічних даних», на *19-й Міжнар. наук.-техн. конф. SAIT 2017*, Київ, 2017, с. 166–167.
- [15] Н. Р. Кондратенко, та О. О. Снігур, «Оцінка перспективності артезіанської свердловини з використанням експертних знань», на *V міжнар. наук.-практ. конф. Інформаційні технології та комп'ютерна інженерія*, Івано-Франківськ, 2015, с. 219-221.
- [16] О. О. Снігур, «Комп'ютерна програма для інтервальної нечіткої класифікації даних», *свідоцтво про реєстрацію авторського права на твір № 63435*, 06.01.2016.

- [17] О. О. Снігур, «Комп'ютерна програма для інтервального нечіткого кластерного аналізу», *свідчення про реєстрацію авторського права на твір № 63436*, 06.01.2016.

## АНОТАЦІЯ

**Снігур О.О. Нечіткі моделі родовищ підземних вод та методи оцінювання їхнього стану за умов невизначеності.** – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня кандидата технічних наук за спеціальністю 01.05.02 – Математичне моделювання та обчислювальні методи. – Вінницький національний технічний університет, Вінниця, 2018.

Дисертацію присвячено підвищенню обґрунтованості оцінок стану родовищ підземних вод шляхом розробки агрегованої моделі родовища підземних вод та методу оцінювання його стану за умов невизначеності на основі інтервальних нечітких множин типу 2. Запропоновано агреговану нечітку модель родовища підземних вод на основі інтервальних нечітких множин типу 2. Адаптовано модель для розв'язання задачі оцінювання перспективності артезіанської свердловини. Розроблена модель може працювати за умов недостатньої поінформованості особи, яка приймає рішення, про значення вхідних змінних, що дозволяє застосовувати її на будь-якому етапі гідрогеологічної розвідки. Розроблено автоматизований метод виділення інформативних ознак, що використовує знання експерта про предметну галузь та накопичені експериментальні дані. Запропоновано агрегуючий критерій, який дає узагальнену інтервальну оцінку значення вихідної змінної на основі множини моделей. Удосконалено метод кластеризації РСМ шляхом введення інтервальних ступенів належності, що дозволяє враховувати та моделювати невизначеності, пов'язані з браком знань про досліджуваний об'єкт. Інтервальна форма методу дає узагальнену оцінку за різними критеріями якості, що дозволяє уникнути помилкових результатів, пов'язаних із внутрішніми особливостями того чи іншого критерію. Розроблено модель на основі кластеризації для роботи з гідрогеологічними даними в задачі оцінювання стану свердловини. Показано, яким чином запропонована модель використовується для аналізу даних поточного спостереження родовищ підземних вод для раннього виявлення потенційно негативних тенденцій у характері зміни якості води. Створено спеціальне програмне забезпечення для оцінювання запасів підземних вод на основі запропонованих моделей та методів.

Результати дисертаційної роботи впроваджено в Правобережній Геологічній Експедиції ДП «Українська геологічна компанія» та в навчальний процес кафедри захисту інформації Вінницького національного технічного університету.

**Ключові слова:** інтервальні нечіткі множини типу 2, нечітка модель, підземні води, невизначеність, кластерний аналіз, інтервальні ступені належності.

**Снигур О. А. Нечеткие модели месторождений подземных вод и методы оценки их состояния в условиях неопределённости.** – Квалификационный научный труд на правах рукописи.

Диссертация на соискание ученой степени кандидата технических наук по специальности 01.05.02 – Математическое моделирование и вычислительные методы. – Винницкий национальный технический университет, Винница, 2018.

Диссертация посвящена повышению обоснованности оценок состояния месторождений подземных вод путём разработки агрегированной модели месторождения подземных вод и метода оценки его состояния в условиях неопределенности на основе интервальных нечетких множеств типа 2. Предложена агрегированная нечеткая модель месторождения подземных вод на основе интервальных нечетких множеств типа 2. Модель адаптирована для решения задачи оценки перспективности артезианской скважины. Разработанная модель может работать в условиях недостаточной информированности лица, принимающего решение, о значениях входных переменных, что позволяет применять ее на любом этапе гидрогеологической разведки. Разработан автоматизированный метод выделения информативных признаков, использующий знания эксперта о предметной области и накопленные экспериментальные данные. Предложен агрегирующий критерий, дающий обобщенную интервальную оценку значения выходной переменной на основе множества моделей. Усовершенствован метод кластеризации РСМ путем введения интервальных степеней принадлежности, что позволяет учитывать и моделировать неопределенности, связанные с недостатком знаний об исследуемом объекте. Интервальная форма метода дает обобщенную оценку по различным критериям качества, что позволяет избежать ошибочных результатов, связанных с внутренними особенностями того или иного критерия. Разработана модель на основе кластеризации для работы с гидрогеологическими данными в задаче оценки состояния скважины. Показано, каким образом предложенная модель используется для анализа данных текущего наблюдения месторождений подземных вод для раннего выявления потенциально отрицательных тенденций в характере изменений качества воды. Создано специальное программное обеспечение для оценки запасов подземных вод на основе предложенных моделей и методов.

Результаты диссертационной работы внедрены в Правобережной Геологической Экспедиции ГП «Українська геологічна компанія» и в учебный процесс кафедры защиты информации Винницкого национального технического университета.

**Ключевые слова:** интервальные нечеткие множества типа 2, нечёткая модель, подземные воды, неопределённость, кластерный анализ, интервальные степени принадлежности.

**Snihur O.O. Fuzzy models of groundwater deposits and methods of evaluating their state under conditions of uncertainty.** – Qualification scientific work published as a manuscript.

Thesis applying for the Ph.D. degree in specialty 01.05.02 – Mathematical Modeling and Computation Methods. – Vinnytsia National Technical University, Vinnytsia, 2018.

The purpose of the thesis research is improving validity of groundwater deposits state

evaluations by developing an aggregated groundwater deposit model and a method for deposit state evaluation under conditions of uncertainty based on interval type-2 fuzzy sets.

A groundwater deposit as an observation object has a number of special features creating uncertainty and making any precise estimations during hydrogeological data analysis impossible. For a mathematical description of natural systems of such complexity it does not usually suffice to state a single analytical function, and it is only possible to judge the way they work based on a data set of experimental data, or on an expert's empirical knowledge. Deposit models are presented with the requirement to adequately reflect the object on every stage of the research, even under a condition of the decision maker being not sufficiently informed about the values of input variables. This causes the importance of developing new mathematical models of groundwater deposits and methods of evaluating their state using interval type-2 fuzzy sets that allow to account for and model the uncertainties present in the input data, and to obtain meaningful results on earlier stages of deposit research.

An aggregated fuzzy groundwater deposit model under conditions of uncertainty based on interval type-2 fuzzy sets is proposed, which has a capability to reflect and model uncertainties, including those caused by the decision maker being not sufficiently informed about part of the input variables' values. Its main functional component is an interval fuzzy logic system with  $m$  input variables  $x_1^z, \dots, x_m^z$  and one output variable  $Y$ . The input vector may contain only partial data about an artesian well, resulting from the decision maker being not sufficiently informed about the values of input variables:  $x_i^z \in \emptyset, i = 1, \dots, k, k < m$ . This is a frequently occurring situation in hydrogeological research, since some of the parameters may not be accessible for direct observation, especially on early stages of the research.

An input vector is presented to the interval fuzzy logic system unmodified, as well as after passing through an informative feature selection procedure. This procedure may be performed by one or more experts in the subject area. In case more than one expert is available, each of them generates their own feature subset, and as a result a separate model. The interval output of the fuzzy logic system obtained on the complete input vector and the outputs of models resulting from discarding uninformative features are combined according to an aggregating criterion, which provides an interval estimation of the output variable value based on a set of models.

As a source of input features for one of the submodels a method of automatic informative feature selection is proposed as an alternative to an expert. It is based on a informative feature search genetic algorithm with clustering.

The developed mathematical model allows for integration of alternative submodels based on other decision making methods or formal procedures that may already exist for solving the given task. In case of its use for the task of artesian well state evaluation the alternative model serves for monitoring changes in a well's state in time by controlling changes in well object position relative to compact groups of training set wells. The model is utilized on the final stage of hydrogeological research and in further monitoring process, and is based on a decision making procedure with a cluster analysis method, namely the modified Possibilistic C-Means (PCM) clustering method with an interval output. The improvement of the PCM method lies in expanding its capabilities by introducing interval membership grades. This allows to account for and model uncertainties related to missing

expert knowledge.

Mathematical description of these uncertainties becomes possible due to presenting the degree of fuzziness in interval form by detecting its margin values according to a number of clustering validity indices. Accounting for and modeling uncertainties inherent in the source data set allows in a number of cases to spot trends and changes in the processes taking place in a groundwater deposit as they originate.

Another advantage of this approach is the ability to work with several input vector simultaneously. Presenting data of several wells to the system's input may not only allow to estimate their position relative to standard compact classes, but also to evaluate their mutual positions.

Experimental testing of the proposed models and methods' functional capabilities in evaluating artesian wells on various stages of hydrogeological research has displayed that the aggregated groundwater deposit fuzzy model allows to receive meaningful results on the final stage, as well as on intermediate stages of geological exploration. In a number of cases the results allow to understand that further operation at the selected site may face certain challenges, and this conclusion may already be reached at the beginning of hydrogeological research. Based on this information, additional research may be conducted and a decision to halt the operations taken, which would allow to spare significant cost on materials and human resources.

Based on the proposed models and methods artesian well prospect evaluation software was developed.

The results of the thesis are implemented at the Pravoberezhna Heolohichna Ekspeditsiya, "Ukrayinska Heolohichna Kompaniya" SE and the learning process of the Information Security Department, Vinnytsia National Technical University.

**Keywords:** interval type-2 fuzzy sets, fuzzy model, groundwater, uncertainty, cluster analysis, interval membership grades.



Підписано до друку 05.09.2018 р. Формат 29,7×42 ¼  
Наклад 100 прим. Зам. № 2018-150.  
Віддруковано в інформаційному редакційно-видавничому центрі  
Вінницького національного технічного університету  
м. Вінниця, Хмельницьке шосе, 95. Тел.: 65-18-06  
Свідоцтво суб'єкта видавничої справи  
серія ДК №3516 від 01.07.2009 р.