
doi

УДК 519.6:504.064

А.В. Яцишин¹, д-р техн. наук, **Ю.Г. Куцан**¹, д-р техн. наук,
В.О. Артемчук¹, канд. техн. наук, **І.П. Каменева**¹, канд. техн. наук,
О.О. Попов¹, д-р техн. наук, **В.О. Ковач**², канд. техн. наук

¹ Інститут проблем моделювання в енергетиці ім. Г. М. Пухова НАН України
(Україна, 03164, Київ, вул. Генерала Наумова 15,
тел. (044) 4249168; e-mail: ak24avo@gmail.com),

² Національний авіаційний університет
(Україна, 03058, Київ, проспект Космонавта Комарова, 1)

Принципи та методи управління екологічною безпекою на основі інтелектуального аналізу даних мережі моніторингу атмосферного повітря*

Проблему управління екологічною безпекою розглянуто щодо зменшення негативних впливів на навколишнє середовище. Запропоновано узагальнену структурну модель процесу управління екологічною безпекою, засновану на методах та технологіях інтелектуального аналізу даних моніторингу. Досліджено можливості адаптації та удосконалення найбільш відомих алгоритмів інтелектуального аналізу даних, таких як C4.5, K-means, методу опорних векторів (SVM), kNN, наївного багатого класифікатора, алгоритму Аргіогі, для задач аналізу даних мережі моніторингу атмосферного повітря. Наведено приклади практичного використання окремих методів з метою виявлення небезпечних ситуацій.

Ключові слова: екологічна безпека, управління безпекою, інтелектуальний аналіз, моніторинг, атмосферне повітря.

Інтенсивний розвиток науково-технічного прогресу в останні десятиріччя супроводжується збільшенням техногенного впливу на атмосферу, гідросферу та літосферу, зростанням обсягів відходів виробничої діяльності й транспортних засобів. Серед глобальних екологічних проблем найбільш пріоритетними є забруднення повітря, дефіцит прісної води, зменшення видового й ландшафтного різноманіття планети, парниковий ефект, озонні діри, кислотні дощі, масові захворювання людей, загибель лісів тощо. Зменшення рівня антропогенного впливу на біосферу можна досягти якіс-

* Роботу виконано в межах проекту, що фінансується за рахунок коштів бюджетної програми «Підтримка розвитку пріоритетних напрямів наукових досліджень» (КПКВК 6541230).

ним управлінням екологічною безпекою на всіх рівнях, забезпечивши стратегічну орієнтацію на принципи стійкого (гармонійного) розвитку.

На даний час існує багато визначень і трактувань щодо поняття «екологічної безпеки». Згідно ст. 50 закону України «Про охорону навколишнього природного середовища» термін «екологічна безпека» визначає такий стан навколишнього природного середовища (ПС), у якому реалізовано реальні можливості для попередження погіршення екологічних умов та виникнення небезпеки для здоров'я людей. Як зазначено в Енциклопедії Сучасної України [esu.com.ua], екологічну безпеку можна розглядати як рівень захищеності життєво важливих інтересів людини, а також суспільства, держави від реальних або потенційних загроз, зумовлених антропогенними або природними чинниками.

Головним критерієм для оцінювання екологічної безпеки певної екосистеми є якість життя і здоров'я населення. Отже, виникає необхідність цілеспрямованого управління екологічною системою з метою підвищення її організованості та досягнення певного ефекту щодо зменшення негативного впливу на населення.

Проблема управління екологічною безпекою. За визначенням американських учених Д. Кліланда і В. Кінга управління можна розглядати як процес, орієнтований на досягнення певних цілей [1]. Для вирішення проблеми необхідно змінити існуючий стан речей (або подій), щоб досягти своєї мети. При всій різноманітності форм їх впливу можна поділити на два протилежні класи: впливи, що призводять до деградації, руйнування екосистеми, зменшення ступеня її стабільності, та впливи, які сприятимуть розвитку екосистеми, збільшенню ступеня її організованості.

Отже, процес управління екологічною безпекою на будь-якому рівні включає певну сукупність послідовних дій, а саме:

- збір інформації;
- передача інформації в пункти збереження та її обробка;
- аналіз збереженої та довідкової інформації;
- прийняття рішення на основі проведеного аналізу;
- створення відповідного керуючого впливу;
- доведення цього впливу до об'єкта управління;
- перевірка нового стану об'єкта управління.

Управління екологічною безпекою та раціональним використанням природних ресурсів передбачає аналіз потреб людини в природних ресурсах і з'ясування можливостей природи щодо задоволення цих потреб. Однак необхідно не тільки виявити оптимальне поєднання потреб людини та можливостей природи, але й забезпечити оптимальні умови для переходу від нинішнього далеко не ідеального стану до більш ефективного варіанту такого поєднання.

На підставі власного досвіду роботи в галузі екологічної безпеки та на основі аналізу наукових джерел [2—7] визначено, що основною метою досліджень в даному напрямі є розробка і впровадження сучасних механізмів управління екологічною безпекою та станом навколишнього середовища. Під словами «механізми управління» будемо розуміти сукупність певних методів та засобів управління екологічною безпекою й природо-користуванням. У свою чергу, метод управління визначає набір способів, прийомів, засобів впливу на керований об'єкт.

У нашій країні прийняття управлінських рішень та реалізація конкретних регулюючих заходів, які необхідно здійснити щодо збереження навколишнього середовища, відбуваються із затримкою. Час затримки $\tau_{ч.з}$ може бути значним і залежить від багатьох факторів. За цей час система може вийти за межі області допустимих станів, які відповідають прийнятному рівню якості навколишнього середовища, тобто нормальним умовам для проживання та діяльності людей на досліджуваній території. Більше того, після досягнення $\tau_{ч.з}$ певного критичного значення екологічна система може перейти в область критичного відхилення, де незначні зміни призведуть до незворотних процесів деградації.

подамо об'єкт управління екологічною безпекою і станом навколишнього середовища у вигляді двох взаємодіючих підсистем:

1) природно-територіальний комплекс, тобто компоненти, на які впливають об'єкти промисловості й транспорту на досліджуваній території, де поширюються та накопичуються забруднюючі речовини;

2) безпосередньо промислові об'єкти і транспорт (ПОТ), які функціонують на досліджуваній території і належать підсистемі стаціонарних джерел забруднення та підсистемі пересувних джерел забруднення.

На рис. 1 наведено узагальнену структурну модель об'єкта управління екологічною безпекою [3], побудовану з урахуванням підходу, викладеного в роботі [4]. На схемі позначено окремі інформаційні потоки, які визначають взаємодію між параметрами: X — вектор стану (якості) компонентів ПС (концентрації забруднень, рівень енергетичних впливів тощо), тобто це контрольовані показники якості компонентів ПС; Z — параметри ПОТ, які визначають умови і результат функціонування окремих об'єктів як з економічного (обсяги, якість продукції та послуг), так і з екологічного погляду (потужність негативного впливу на ПС).

Для даної моделі $Z = \{Z_c, Z_n\}$, де Z_c, Z_n — множини станів відповідно стаціонарних та пересувних об'єктів ПОТ. При цьому x_c, x_n — множини впливів ПС на підсистеми ПОТ ($x_c, x_n \subset X$); z'_c, z'_n — множини результатів економічної діяльності об'єктів ПОТ ($z'_c \subset Z_c, z'_n \subset Z_n$); $\omega_{n,c}$ — множина зовнішніх впливів на компоненти ПС. У такому вигляді можна

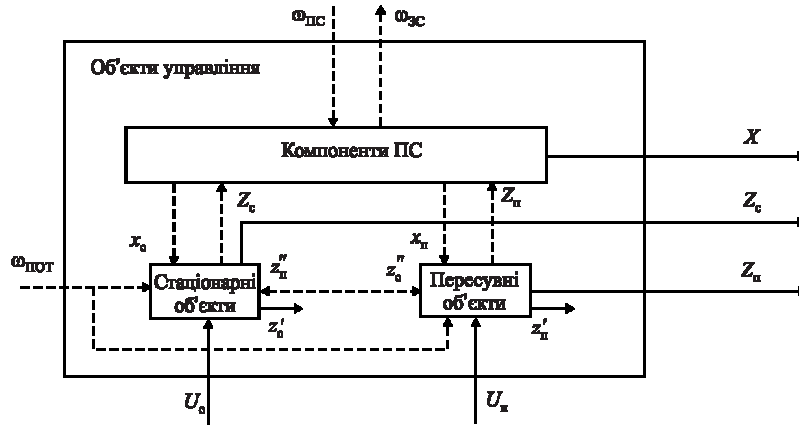


Рис. 1. Узагальнена модель управління екологічною безпекою територій, що зазнають впливу ПОТ

врахувати природні та кліматичні параметри, особливості інфраструктури регіону тощо. Враховано також додаткові зовнішні впливи, пов'язані з ПОТ, а саме: $\omega_{\text{ПОТ}}$ — множина зовнішніх впливів на ПОТ, які впливають на результати їх діяльності, виробничі й транспортні параметри, що визначають рівень екологічної небезпеки; $\omega_{\text{ЗС}}$ — множина впливів на зовнішнє середовище (ЗС) через компоненти ПС, що надходять від ПОТ; $U_c, U_{\text{п}}$ — множина керуючих впливів відповідно на стаціонарні та пересувні ПОТ.

Особам, що приймають рішення (ОПР) щодо виникнення в момент t несприятливої екологічної ситуації, необхідно завчасно надати інформацію для розробки керуючих впливів $U_c(t)$ або $U_{\text{п}}(t)$, за допомогою яких можна контролювати та зменшувати негативний вплив на компоненти ПС: $\Delta X(t) < \varepsilon, \varepsilon \rightarrow 0$.

Розглянемо детальніше поняття стратегічного і тактичного управління екологічною безпекою, оскільки вони мають вирішальне значення в процесах управління. Стратегічне і тактичне управління екологічною безпекою мають різні цілі та засоби їх досягнення. Якщо стратегічне управління являє собою пошук і реалізацію нових можливостей при змінах ЗС, то тактичне управління є процесом створення передумов для реалізації певних нових можливостей.

Тактичне управління можна визначити як засіб реалізації та деталізації відповідної стратегії. Воно розраховано на короткостроковий та середньостроковий періоди. Головною метою тактичного управління екологічною безпекою є вплив на життєдіяльність досліджуваної екосистеми, спрямований на досягнення стратегічних цілей та завдань при найбільш повному і найраціональнішому використанні доступних ресурсів. Зреш-

тою, засобом реалізації стратегічних планів є тактичне планування. Якщо основна мета стратегічного плану полягає у визначенні перспективи на майбутнє, то тактичне планування допомагає знайти конкретні відповіді на питання, як поетапно досягти бажаного стану.

На рис. 2 наведено концептуальну схему системи управління екологічною безпекою урбанізованої території за даними екологічного моніторингу з використанням сучасних методів і технологій інтелектуального аналізу даних [3]. Дано пояснення основних блоків, зазначених на схемі.

Блок «Бази даних та алгоритми оцінювання ситуації і ухвалення управлінських рішень в галузі екологічної безпеки» складається з блоків «Управління даними» та «Система підготовки рішень (на основі економічної оцінки ризик-ціна-ефект)». У свою чергу, блок «Система підготовки рішень» включає «Інтелектуальний аналіз даних» і «Систему підтримки прийняття рішень (СППР)».

Блок «Управління даними» забезпечує збереження даних і доступ до них. До таких даних відносять дані моніторингу, розподілені бази даних (БД) аналітичної та звітної інформації та розподілені БД дистанційного зондування Землі (ДЗЗ), які перебувають у сховищі даних. Під розподіленими БД розуміємо сукупність взаємопов'язаних на логічному рівні БД, розподілених у комп'ютерній мережі. Інтелектуальний аналіз даних у наведеній схемі включає експертні системи, геоінформаційні та інтеграційні технології (ДЗЗ, Інтернет).

В цілому інтелектуальний аналіз даних у задачах екологічної безпеки можна визначити як метод підтримки прийняття рішень, заснований на дослідженні певних закономірностей щодо вихідних даних. Він включає методики та засоби, які на основі будь-яких моделей, алгоритмів або математичних теорем дають можливість оцінювати значення невідомих характеристик і параметрів за відомими даними.

Алгоритми інтелектуального аналізу. На основі публікацій [8—17, 21] визначено найпоширеніші алгоритми інтелектуального аналізу даних, спрямовані на виявлення прихованих закономірностей або взаємозв'язків між змінними у великих масивах необроблених даних.

Класифікатор C4.5 створює класифікацію у вигляді дерева рішень. Для цього задається набір даних, який являє собою вже класифіковані речі. Алгоритм C4.5, розроблений J.R. Quinlan [13], є вдосконаленою версією алгоритму ID3 того ж автора. Зокрема, в C4.5 додано відсікання гілок, можливість роботи з числовими атрибутами, а також можливість побудови дерева з неповною навчальною вибіркою, в якій відсутні значення деяких атрибутів.

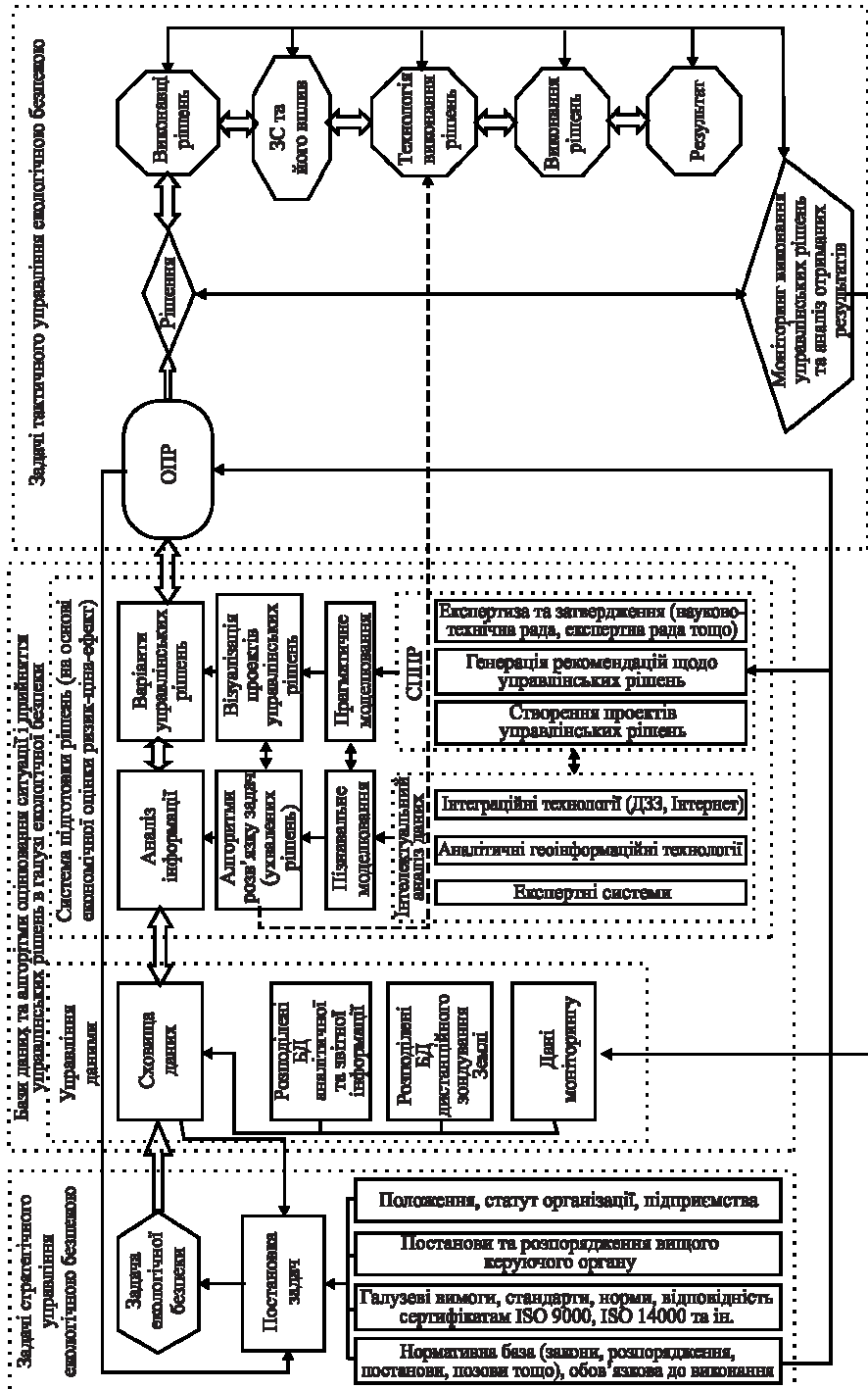


Рис. 2. Концептуальна схема системи управління екологічною безпекою за даними екологічного моніторингу

Для того щоб за допомогою С4.5 побудувати дерево рішень і застосувати його, дані повинні задовольняти декільком умовам, а саме:

інформація про об'єкти, що необхідно класифікувати, повинна бути надана у вигляді кінцевого набору ознак (атрибутів), кожен з яких має дискретне або числове значення. Якщо такий набір атрибутів назвати прикладом, тоді для всіх прикладів кількість атрибутів і їх склад повинні бути постійними;

кожен із класів, на які будуть розбиватися приклади, повинен мати кінцеве число елементів, а кожен приклад повинен однозначно відноситись до конкретного класу;

для випадків з нечіткою логікою, тобто коли приклади можна віднести до одного з класів з певною ймовірністю, С4.5 непридатний;

у навчальній вибірці кількість прикладів має бути значно більше кількості класів і кожен приклад повинен бути заздалегідь асоційований зі своїм класом.

За цими ознаками С4.5 віднесено до методів машинного навчання з учителем.

Алгоритм кластерного аналізу K-means виділяє певну кількість груп k з набору об'єктів таким чином, щоб члени кожної групи були максимально схожими між собою. Це досить популярний метод кластерного аналізу для вивчення набору даних. Нагадаємо, що кластерний аналіз — це сімейство алгоритмів, призначених для формування груп, де в кожній групі члени даної групи мають більше схожих рис з представниками своєї групи, ніж з тими, хто в цій групі не перебуває (в даному контексті кластери та групи є синонімами) [15].

Метод K-means — це метод, метою якого є поділ m спостережень (з простору даних) на k кластерів таким чином, що кожне спостереження буде віднесено до того кластеру, до центру (центроїду) якого воно буде найближчим. Як міра близькості використовується евклідова відстань:

$$\|x - y\| = \sqrt{\sum_{p=1}^n (x_p - y_p)^2}, \quad x, y \in R^n.$$

Для ряду спостережень (x^1, x^2, \dots, x^m) , $x_j \in R^n$, метод K-means мінімізує сумарне квадратичне відхилення точок кластерів від центроїдів цих кластерів:

$$\min \left[\sum_{i=1}^k \sum_{x^{(j)} \in S_i} \|x^{(j)} - \mu_i\|^2 \right],$$

де $x^{(j)} \in R^n$; $\mu_i \in R^n$; μ_i — центр ваги для кластера S_i .

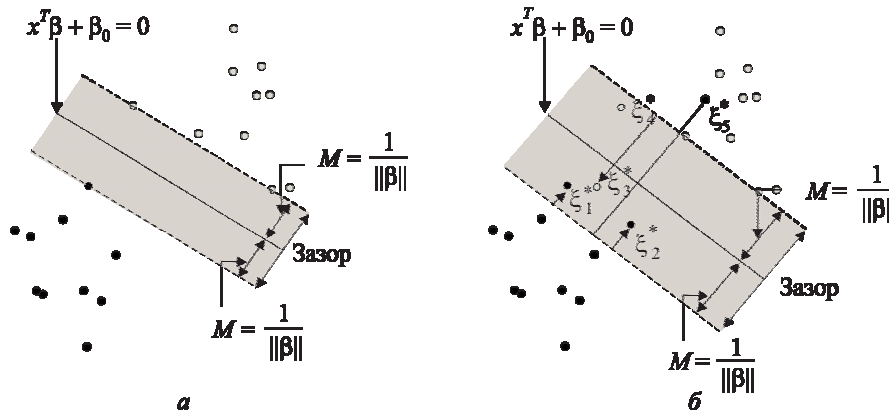


Рис. 3. Класифікатори з мінімальним зазором (а) і на опорних векторах (б)

Зауважимо, що при використанні даного методу неправильний вибір початкового числа кластерів k може призвести до некоректних результатів. Саме тому при використанні методу K-means на початковому етапі необхідно визначити відповідне число кластерів для визначеного набору даних.

Метод опорних векторів (SVM) дозволяє знаходити гіперплощину для розподілу даних на два класи. На відміну від C4.5 в ньому не використовуються дерева рішень. SVM — це лінійний алгоритм, який використовується в задачах класифікації та регресії. Він широко застосовується на практиці і за його допомогою можна вирішувати як лінійні, так і нелінійні задачі. Суть цього методу полягає у створенні лінії або гіперплощини, яка розділяє дані на класи. Якщо є два класи спостережень і передбачається лінійна форма кордону між класами, то можливі два варіанти.

Перший з них пов'язаний з можливістю ідеального поділу даних за допомогою деякої гіперплощини $z_k(x) = \sum_{i=1}^p \beta_i x_i + \beta_0$. На рис. 3, а, наведено двовимірний варіант. Оскільки таких гіперплощин може бути безліч, оптимальною буде та з них, яка максимально віддалена від навчальних точок, тобто має максимальний проміжок (зазор) M [16].

Другий варіант показано на рис. 3, б, коли множина точок перекривається і обидва класи лінійно нероздільні. Власне опорними векторами називаються значення, які лежать безпосередньо на кордоні, що розділяє смуги, або на неправильній для свого класу стороні щодо кордонів зазору (такі точки позначені ξ_j^*). Для граничних і всіх інших точок прийнято $\xi_j^* = 0$.

Оптимальну розділяючу гіперплощину такого класифікатора $zk(x)$ можна визначити за умови максимізації ширини зазору M , але при цьому дозволено невірно класифікувати деяку невелику групу спостережень, що відносяться до опорних векторів. Математично пошук рішення зведено до задачі квадратичної оптимізації з лінійними обмеженнями, яка гарантовано сходиться до одного глобального мінімуму.

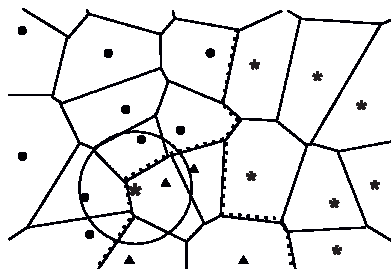


Рис. 4. Приклад роботи алгоритму класифікації kNN найближчих сусідів [16]

Алгоритм класифікації kNN (k найближчих сусідів) відрізняється від раніше описаних тим, що він нагадує «лінивого учня». В основі методу kNN-класифікатора лежить гіпотеза компактності, яка передбачає, що тестований об'єкт d матиме таку ж мітку класу, як і навчальні об'єкти в локальній області його найближчого оточення. У варіанті kNN кожен об'єкт відносимо до пріоритетного класу найближчих сусідів, де k — параметр алгоритму.

Вирішальні правила в методі kNN визначаються межами суміжних сегментів діаграми Вороного, що розділяє площину на n опуклих багатокутників, кожен з яких містить один і тільки один об'єкт навчальної вибірки (рис. 4). В p -мірних просторах границі розв'язків складаються вже з сегментів $(p - 1)$ -мірних напівплощин, утворених опуклими многогранниками Вороного [16].

Алгоритм передбачення будується за принципом «більшості голосів», тобто за результатами голосування вибирається мітка класу-переможця. На рис. 4 тестований об'єкт «зірочка» потрапляє в осередок об'єктів класу «трикутників» і при $k = 1$ буде віднесений до цього класу. Однак при $k = 3$ за «голосами» двох найближчих сусідів з трьох екзаменованих цей об'єкт буде віднесено до класу «кружечків». Імовірнісний варіант методу kNN використовує для ранжирування передбачуваних класів суму голосів сусідів з урахуванням їх ваг, зокрема евклідової міри відстані між тестованим об'єктом і кожним сусідом.

Варіант 1NN завжди забезпечує 100% правильного розпізнавання прикладів навчальної вибірки (самий найближчий сусід — він сам), проте часто помиляється на невідомих йому даних. При збільшенні k від одиниці до деяких меж якість розпізнавання на контрольній вибірці буде зростати. Оптимальне щодо точності прогнозів значення k може бути знайдено з використанням перехресної перевірки. Для цього по фіксованому значенню k будується модель k найближчих сусідів і оцінюється CV-помилка

класифікації. Ці дії повторюються для різних значень k і значення з найменшою помилкою розпізнавання приймається як оптимальне.

Наївний байєсів класифікатор (Naive Bayes Classifier) включає сімейство алгоритмів класифікації, які поділяють одне спільне припущення. В основі байєсівської класифікації лежить гіпотеза максимальної ймовірності, тобто об'єкт d належить класу c_j ($c_j \in C$), і при цьому досягається найбільша апостеріорна ймовірність $\max P(c_j|d)$. За формулою Байєса

$$P(c_j|d) = \frac{P(c_j)P(d|C_j)}{P(d)} \approx P(c_j)P(d|c_j),$$

де $P(d|c_j)$ — ймовірність зустріти об'єкт d серед об'єктів класу c_j ; $P(c_j)$ і $P(d)$ — апіорні ймовірності класу c_j і об'єкта d (остання не впливає на вибір класу і її можна не враховувати).

Якщо зробити «наївне» припущення, що всі ознаки, за якими класифікуються об'єкти, абсолютно рівноправні між собою і не пов'язані одна з одною, то $P(d|c_j)$ можна обчислити як добуток ймовірностей зустріти ознаку x_i ($x_i \in X$) серед об'єктів класу c_j :

$$P(d|c_j) = \prod_{i=1}^{|X|} P(x_i|c_j),$$

де $P(x_i|c_j)$ — імовірнісна оцінка вкладу ознаки x_i в те, що $d \in c_j$.

На практиці при перемноженні дуже малих умовних ймовірностей може спостерігатися втрата значущих розрядів. У зв'язку з цим замість оцінок ймовірностей $P(x_i|c_j)$ застосовують логарифми цих ймовірностей. Оскільки логарифм є монотонно зростаючою функцією, клас c_j з найбільшим значенням логарифма ймовірності залишиться найбільш імовірним. Тоді вирішальне правило наївного байєсового класифікатора приймає наступний остаточний вигляд [16]:

$$c^* = \arg_{c_j \in C} \max \left[\log P(c_j) + \sum_{i=1}^X P(x_i|c_j) \right].$$

Алгоритм Apriori належить до масштабованих алгоритмів, спрямованих на пошук асоціативних правил в БД. Сучасні БД мають дуже великі розміри, а також тенденцію щодо подальшого їх зростання. Пошук асоціативних правил в таких БД спрямований на отримання даних для вивчення кореляцій і взаємозв'язків між змінними. Алгоритм Apriori реалізує такий пошук в два етапи. На першому етапі здійснюється пошук вибірок, які часто зустрічаються в даній базі, а на другому — виявляються асоціативні правила, що є достовірними для цих вибірок.

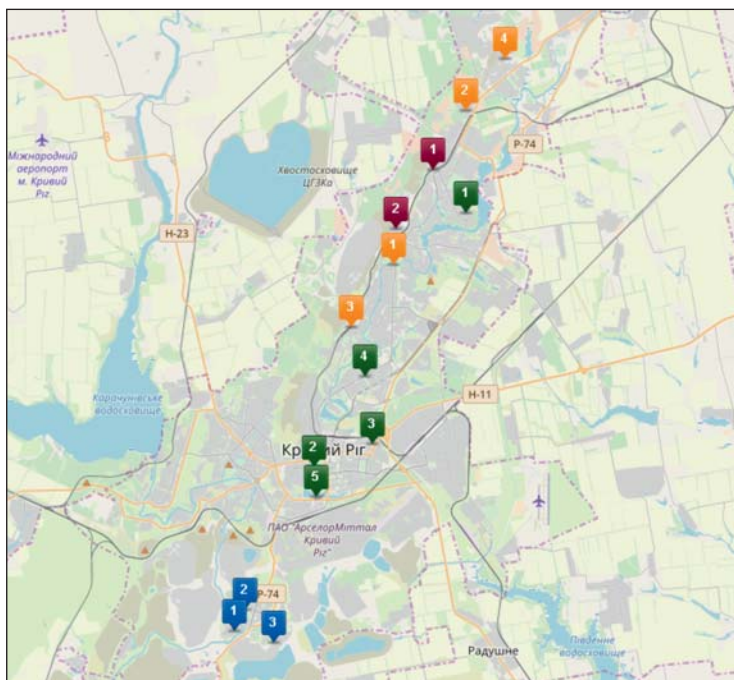
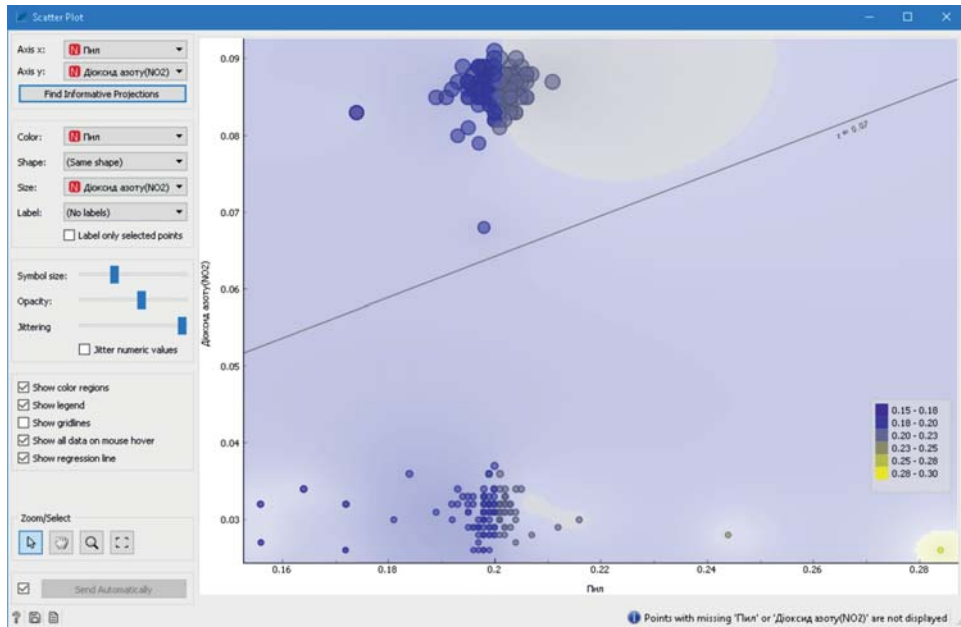


Рис. 5. Автоматизовані пости спостереження за станом АП

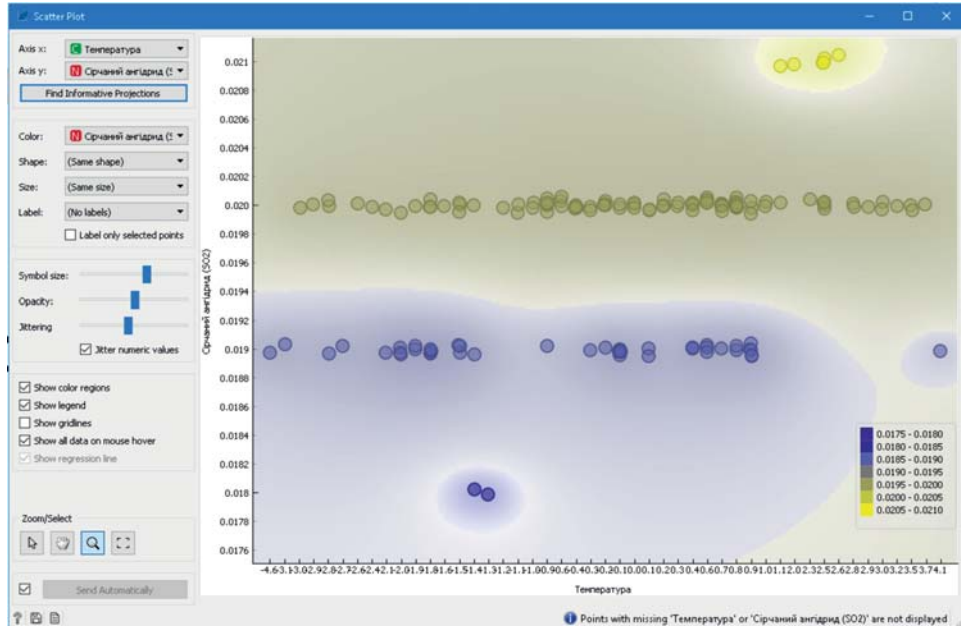
Приклади інтелектуального аналізу даних. Інтелектуальний аналіз даних моніторингу стану атмосферного повітря (АП) урбанізованих територій проведено в рамках інформаційно-аналітичної системи еколого-енергетичного моніторингу AISEEM [3, 18—20] та програми Orange [22], яка є сучасним інструментом для візуалізації даних, машинного навчання та інтелектуального аналізу даних. Взаємодія між цими програмами здійснювалась за допомогою інтеграції методів та засобів Orange в систему AISEEM.

Дані щодо забруднюючих речовин завантажувались з автоматизованої системи моніторингу за станом АП в м. Кривий Ріг (Дніпропетровська область). Основною метою системи моніторингу є безперервні виміри концентрацій забруднюючих речовин і метеорологічних параметрів АП жилої зони м. Кривий Ріг [23].

На рис. 5 темно-зеленим кольором позначені міські автоматизовані пости, темно-синім — пости підприємства ПАТ «Південний гірничо-збагачувальний комбінат», темно-червоним — ПРАТ «СУХА БАЛКА» та помаранчевим — ПАТ «Кривбасзалізрудком». Спостереження здійснено за шістьма забруднюючими речовинами: діоксид азоту (NO_2), оксид азоту



а



б

Рис. 6. Діаграми розсіювання концентрацій пилу і діоксиду азоту (а) та сірчаного ангідриду і температури повітря (б)

(NO), сірчаний ангідрид (SO₂), оксид вуглецю (CO), аміак (NH₃), сірководень (H₂S), пил. Крім того, на міських автоматизованих постах спостережень (ПАС) № 3 та № 5 здійснено спостереження за етаном (C₂H₆) та озоном (O₃), а метеорологічними параметрами є середня швидкість вітру, напрямок вітру, температура, відносна вологість та тиск.

Варто зазначити, що на території міста виробництво теплової енергії здійснює ПАТ «Криворізька теплоцентраль», на балансі якої є шість опалювальних котельень, обладнаних 31 паровими та водогрійними котлами з установленою та наявною потужностями відповідно 1406 Гкал/год та 1081 Гкал/год. Найшкідливіші викиди цього енергетичного об'єкта — це сірчаний ангідрид, оксиди азоту та пил.

При аналізі даних було виключено інформацію щодо напрямку вітру, оскільки напрямок вітру змінюється від 0 до 360°, і для 358° та 2° фактично отримуємо північний напрямок вітру згідно з даними [23] з зовсім різними значеннями, що ускладнює їх подальше трактування.

Наведемо приклади інтелектуального аналізу даних для погодинних максимальних разових концентрацій забруднюючих речовин та метеорологічних даних на міському ПАС № 1 з 10 по 16 грудня 2018 року.

На рис. 6, а, зображено діаграму розсіювання концентрацій пилу та діоксиду азоту. Як бачимо, концентрації пилу на цьому часовому інтервалі змінювались від 0,156 до 0,284 мг/м³, а концентрації діоксиду азоту — від 0,027 до 0,091 мг/м³. На рис. 6, б, показано діаграму розсіювання концентрації сірчаного ангідриду та температуру повітря (°C). Температура на цьому часовому інтервалі змінювалась від -4,1 до +4,6°C, а концентрації сірчаного ангідриду — від 0,018 до 0,021 мг/м³.

Слід зазначити, що на даному етапі дослідження наведено лише попередні результати аналізу даних моніторингу. Адже застосування інтелектуальних технологій для виявлення нових знань та закономірностей потребує значно більшої кількості даних щодо забруднення атмосфери, викидів транспорту та промислових підприємств, захворюваності населення тощо. Маємо надію, що такі дані будуть доступні після реалізації Загальнодержавної автоматизованої системи «Відкрите довкілля», що дасть можливість здійснювати більш ґрунтовний аналіз даних моніторингу стану АП.

Висновки

Запропонована концептуальна модель управління екологічною безпекою урбанізованих територій за даними екологічного моніторингу відрізняється від своїх аналогів новими можливостями, базованими на використанні сучасних інформаційних технологій інтелектуального аналізу даних. Визначено перспективи застосування методів та засобів інтелектуального аналізу даних щодо інформаційної підтримки прийняття рішень, спрямо-

ваних на оцінювання наслідків техногенного впливу і зменшення обсягу навантажень на довкілля. Запропоновані методи та алгоритми дають можливість оцінювати значення невідомих характеристик і параметрів за відомими даними.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. *Кліланд Д., Кінг В.* Системный анализ и целевое управление. М.: Сов. радио, 1974, 280 с.
2. *Пампура В.И.* Оптимальное управление безопасностью экологически опасных объектов. Київ: Наук. думка, 2012.
3. *Яцишин А.В.* Комплексне оцінювання та управління екологічною безпекою при забрудненнях атмосферного повітря. Дисертація ... докт. тех. наук. Київ, 2013, 402 с.
4. *Иващук О.А., Константинов И.С.* Обеспечение адаптивного управления экологической безопасностью промышленно-транспортного комплекса // Управление большими системами, 2009, вып. 25, с. 96—115.
5. *Капралов Е.Г. и др.* Геоинформатика: учебник для вузов. М.: Академия, 2008.
6. *Артемчук В.А., Каменева И.П., Яцишин А.В.* Модели представления и преобразования данных в задачах экологического мониторинга урбанизированных территорий // Электрон. моделирование, 2016, **38**, № 2, с. 49—66.
7. *Артемчук В.А., Каменева И.П., Яцишин А.В.* Специфика применения когнитивного анализа информации в задачах обеспечения экологической безопасности // Там же, 2017, **39**, № 6, с. 107—124.
8. *Chen C., Chuang C., Jiang J.* Ecological Monitoring Using Wireless Sensor Networks—Overview, Challenges, and Opportunities // Advancement in Sensing Technology. Smart Sensors, Measurement and Instrumentation, 2013, vol. 1., p. 1—21.
9. *Dias D. and Tchepel O.* Modelling of human exposure to air pollution in the urban environment: a GPS-based approach // Environmental Science and Pollution Research, 2014, 21.5, p. 3558-3571.
10. *Meier F. et al.* Challenges and benefits from crowdsourced atmospheric data for urban climate research using Berlin, Germany, as testbed // Proc. of the 9th International Conference on Urban Climate, 2015.
11. *Peters D. et al.* Harnessing the power of big data: infusing the scientific method with machine learning to transform ecology // Ecosphere, 2014, 5.6, p. 1-15.
12. *Алгоритмы интеллектуального анализа данных.* Дата доступу 20.12.2017. Режим доступу: <https://proger.ru/translations/top-10-data-mining-algorithms/>
13. *Деревья решений — С4.5 математический аппарат. Часть 1* [Электронный ресурс] Веб-сайт. Дата доступу 13.11.2018. Режим доступу : <https://basegroup.ru/community/articles/math-c45-part1>. Загол. з екрану.
14. *Згуровский М.З., Болдак А.А., Ефремов К.В.* Интеллектуальный анализ и системное согласование научных данных в междисциплинарных исследованиях // Кибернетика и системный анализ, 2013, 49, № 4, с. 62—75.
15. *Кластеризация: метод k-средних. Часть 1* [Электронный ресурс] Веб-сайт. Дата доступу 13.11.2018. Режим доступу : <http://statistica.ru/theory/klasterizatsiya-metod-k-srednikh/>. Загол. з екрану.
16. *Шитиков В.К., Маслицкий С.Э.* Классификация, регрессия, алгоритмы Data Mining с использованием R. [Электронный ресурс] Веб-сайт. Дата доступу 13.11.2018. Режим доступу : <https://github.com/ranalytics/data-mining>. Загол. з екрану.
17. *Артемчук В.О., Яцишин А.В.* Интеллектуальный анализ данных в системе мониторингу стану атмосферного повітря // Моделирование та інформаційні технології, 2018, вип. 82, с. 48—52.

18. *Артемчук В.О., Білан Т.Р., Білінов І.В. та ін.* Теоретичні та прикладні основи економічного, екологічного та технологічного функціонування об'єктів енергетики. Київ, ТОВ «Наш формат», 2017.
19. *Каменева І.П., Попов О.О., Яцишин А.В., Артемчук В.О.* Методи визначення екологічного ризику за атмосферним фактором // *Моделювання та інформаційні технології*, 2009, вип. 53, с. 23—32.
20. *Яцишин А.В., Каменева І.П., Артемчук В.А., Попов А.А.* Методы и технологии анализа рисков для здоровья на основе данных мониторинга. IV Международная научная конференция «МОДЕЛИРОВАНИЕ-2012», Киев, 2012, с. 470—473.
21. *Ashouri, Milad, et al.* Development of building energy saving advisory: A data mining approach // *Energy and Buildings*, 2018, 172, p. 139—151.
22. *Orange* [Електронний ресурс] Веб-сайт. Дата доступу 02.02.2019. Режим доступу : <http://orange.biolab.si/download/>. Загол. з екрану.
23. *Автоматизовані пости спостереження* [Електронний ресурс] Веб-сайт. Дата доступу 02.02.2019. Режим доступу: <https://krmisto.gov.ua/ua/rc/ecomon.html>. Загол. з екрану.

Отримано 24.04.19

REFERENCES

1. Pampuro, V.I. (2012), *Optimal'noye upravleniye bezopasnost'yu ekologicheskii opasnykh ob'yektov* [Optimal safety management of environmentally hazardous facilities], Naukova Dumka, Kiev, Ukraine.
2. Cleland, D. and King, V. (1974), *Sistemnyy analiz i tselevoye upravleniye* [System Analysis and Target Management], Sov. Radio, Moscow, SSSR.
3. Yatsyshyn, A.V. (2013), “Comprehensive assessment and management of environmental safety in air pollution”, Abstract of Doct. Sci. (Tech.) dissertation, 21.06.01, SI «Institute of Environmental Geochemistry of NAS of Ukraine», Kyiv, Ukraine.
4. Ivashhuk, O.A. and Konstantinov I.S. (2009), “Guarding of adaptive management by ecological safety of an industrial and transport complex”, *Upravleniye bol'shimi sistemami*, no. 25, pp. 96-115.
5. Kapralov, E.G. (2008), *Geoinformatika: uchebnyk dlya vuzov* [Geoinformatics: a textbook for universities], Akademiya, Moscow, Russia.
6. Kameneva, I.P., Artemchuk, V.O. and Yatsyshyn, A.V. (2016), “Models of representation and data transformation in the problems of environmental monitoring in urban areas”, *Elektronnoye modelirovaniye*, Vol. 38, no. 2, pp. 49-66.
7. Artemchuk, V.O. Kameneva, I.P. and Yatsyshyn, A.V. (2017), “Specificity of the application of cognitive analysis of information in the tasks of ensuring environmental safety”, *Elektronnoye modelirovaniye*, Vol. 39, no. 6, pp. 107-124.
8. Chen, C., Chuang, C. and Jiang, J. (2013), «Ecological Monitoring Using Wireless Sensor Networks-Overview, Challenges, and Opportunities», *Advancement in Sensing Technology. Smart Sensors, Measurement and Instrumentation*, Vol. 1, pp. 1-21.
9. Dias, D. and Tchepel, O. (2014), “Modelling of human exposure to air pollution in the urban environment: a GPS-based approach”, *Environmental Science and Pollution Research* 21.5, pp. 3558-3571.
10. Meier, F. et al. (2015), “Challenges and benefits from crowdsourced atmospheric data for urban climate research using Berlin, Germany, as testbed”, *Proceedings of the 9th International Conference on Urban Climate*.
11. Peters, D. et al. (2014), “Harnessing the power of big data: infusing the scientific method with machine learning to transform ecology”, *Ecosphere*, Vol. 5, no. 6, pp. 1-15.
12. “Data Mining Algorithms”, available at: <https://tproger.ru/translations/top-10-data-mining-algorithms/> (accessed June 23, 2019).

13. "Decision trees", available at: <https://basegroup.ru/community/articles/math-c45-part1> (accessed June 23, 2019).
14. Zgurovsky, M.Z., Boldak, A.A. and Yefremov, K.V. (2013), "Intelligent analysis and systemic adjustment of scientific data in interdisciplinary research", *Kibernetika i sistemnyy analiz*, Vol. 49, no. 4, pp. 62-75.
15. "Clustering: k-means method", available at: <http://statistica.ru/theory/klasterizatsiya-metod-k-srednikh/> (accessed Mart 23, 2019).
16. Shitikov, V.K., Mastitsky, S.E. "Classification, regression, Data Mining algorithms using R", available at: <https://github.com/ranalytics/data-mining> (accessed February 2, 2019).
17. Artemchuk, V.O. and Yatsyshyn, A.V. (2018), "Intelligent analysis of data in the system of monitoring of atmospheric air", *Modelyuvannya ta Informatsiyini tehnologiyi*, Vol. 82, pp. 48-52.
18. Artemchuk, V.O., Bilan, T.R., Blinov, I.V. et al. (2017), Theoretical and applied bases of economic, ecological and technological functioning of energy objects. Kyiv, Ukraine: TOV «Nash format», (in Ukrainian).
19. Kameneva, I.P., Popov, O.O., Yatsyshyn, A.V. and Artemchuk, V.O. (2009), "Methods for determining environmental risk factor for atmospheric", *Modelyuvannya ta Informatsiyini tehnologiyi*, Vol. 53, pp. 23-32.
20. Yatsyshyn, A.V., Kameneva, I.P., Artemchuk, V.A. and Popov, A.A. (2012), "Methods and technologies for risk analysis for health based on monitoring data", *Conference proceedings of the 4 th International scientific conference "MODELING-2012"*, 2012, Kyiv, Ukraine, pp. 470-473.
21. Ashouri, Milad et al. (2018), "Development of building energy saving advisory: A data mining approach", *Energy and Buildings*, Vol. 172, pp. 139-151.
22. Orange, available at: <http://orange.biolab.si/download/> (accessed February 2, 2019).
23. "Automated surveillance posts", available at: <https://krmisto.gov.ua/ua/rc/ecomon.html> (accessed February 2, 2019).

Received 24.04.19

А.В. Яцишин, Ю.Г. Куцан, В.А. Артемчук,
И.П. Каменева, А.А. Попов, В.Е. Ковач

ПРИНЦИПЫ И МЕТОДЫ УПРАВЛЕНИЯ ЭКОЛОГИЧЕСКОЙ БЕЗОПАСНОСТЬЮ НА ОСНОВЕ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ СЕТИ МОНИТОРИНГА АТМОСФЕРНОГО ВОЗДУХА

Проблема управления экологической безопасностью рассмотрена относительно уменьшения негативных воздействий на окружающую среду. Предложена обобщенная структурная модель процесса управления экологической безопасностью, основанная на методах и технологиях интеллектуального анализа данных мониторинга. Исследованы возможности адаптации и совершенствования наиболее известных алгоритмов интеллектуального анализа данных, таких как C4.5, K-means, метода опорных векторов (SVM), kNN, наивного байесового классификатора, алгоритма Apriori, для задач анализа данных сети мониторинга атмосферного воздуха. Приведены примеры практического использования отдельных методов с целью выявления опасных ситуаций.

К л ю ч е в ы е с л о в а: экологическая безопасность, управление безопасностью, интеллектуальный анализ, данные мониторинга, атмосферный воздух.

*A.V. Iatsyshyn, Yu. G. Kutsan, V.O. Artemchuk,
I.P. Kameneva, O.O. Popov, V.O. Kovach*

THE PRINCIPLES AND METHODS OF ECOLOGICAL SAFETY MANAGEMENT
THROUGH THE DATA OF AIR MONITORING NETWORK ANALYSIS

The problem of environmental safety management is considered in the context of reducing negative environmental impacts. A generalized structural model of ecological safety management process is proposed, which is based on the methods and technologies of the intellectual analysis of monitoring data. The possibilities of adaptation and improvement of some of the most well-known algorithms for data mining: C4.5, K-means, SVM, kNN, naive Bayes classifier, Apriori algorithm for data analysis of atmospheric air monitoring network data were explored. Examples of practical use of separate methods for the detection of dangerous situations are given.

Key words: ecological safety, security management, intelligence analysis, monitoring data, atmospheric air.

ЯЦИШИН Андрій Васильович, д-р техн. наук, в.о. пров. наук. співр. Ін-ту проблем моделювання в енергетиці ім. Г.С. Пухова НАН України. У 2002 р. закінчив Київський національний університет ім. Т.Г. Шевченка. Область наукових досліджень — математичне моделювання екологічних процесів, екологічний моніторинг техногенних об'єктів, інформаційні технології.

КУЦАН Юлій Григорович, д-р техн. наук, в.о. заст. директора з наукової роботи Ін-ту проблем моделювання в енергетиці ім. Г.С. Пухова НАН України. У 1966 р. закінчив Київський політехнічний ін-т. Область наукових досліджень — моделювання технологічних процесів в енергетичній галузі.

АРТЕМЧУК Володимир Олександрович, канд. техн. наук, ст. наук. співр. Ін-ту проблем моделювання в енергетиці ім. Г.С. Пухова НАН України. У 2008 р. закінчив Житомирський державний технологічний університет. Область наукових досліджень — математичне моделювання та чисельні методи, інформаційні технології.

КАМЕНЕВА Ірина Петрівна, канд. техн. наук, ст. наук. співр. Ін-ту проблем моделювання в енергетиці ім. Г.С. Пухова НАН України. У 1976 р. закінчила Київський державний університет ім. Тараса Шевченка. Область наукових досліджень — аналіз даних і математичне моделювання в екології.

ПОПОВ Олександр Олександрович, д-р техн. наук, ст. наук. співр., пров. наук. співр. Ін-ту проблем моделювання в енергетиці ім. Г.С. Пухова НАН України. У 2004 р. закінчив Житомирський державний технологічний університет. Область наукових досліджень — математичне моделювання забруднення довкілля, вирішення актуальних задач цивільного захисту територій, навколишнього природного середовища та населення.

КОВАЧ Валерія Омелянівна, канд. техн. наук, заст. директора Навчально-наукового інституту неперервної освіти Національного авіаційного університету. У 2011 р. закінчила Національний авіаційний університет. Область наукових досліджень — впровадження наукових засад освіти впродовж життя, державні механізми управління ринком праці.