

# ПРИКЛАДНА ЯДЕРНА ГАММА-СПЕКТРОМЕТРІЯ В ДЕЯКИХ ЗАДАЧАХ ІДЕНТИФІКАЦІЇ КЕРАМІЧНИХ ВИРОБІВ

О.М. Матьовка<sup>1</sup>, М.В. Стець<sup>2</sup>, В.Т. Маслюк<sup>2</sup>, В.М. Бузаш<sup>1</sup>

<sup>1</sup> Ужгородський національний університет, м. Ужгород

<sup>2</sup> Інститут електронної фізики НАН України, м. Ужгород,  
e-mail: an@zvl.iep.uzhgorod.ua

На основі значень питомих активностей (Ап)  $^{40}\text{K}$ ,  $^{228}\text{Ac}$  (ряд  $^{232}\text{Th}$ ),  $^{214}\text{Pb}$  (ряд  $^{238}\text{U}$ ) зразків доквілля (грунтів, глини та керамічних виробів з них) здійснена оцінка можливості їх використання в різних методах статистичної обробки: аналізів розподілів частотей, варіаційних рядів, діаграм “Ап-Ап-к, Ап/Ап-Ап/Ап, матриць дистанцій, для вирішення задач ідентифікації. Взаємоузгодженість результатів цих методів підтверджує правильність факторизації множин значень Ап для різних типів зразків.

## Вступ

Ми розглядаємо можливості прикладної ядерної гамма-спектрометрії (ПЯГС) на основі отриманого масиву експериментальних даних декількох типів зразків керамічних виробів, глини та ґрунтів. Керамічні вироби вибрані тому, що вони є суттєвим матеріальним та радіаційним компонентом нашого доквілля і об'єктами, на яких можна виявити вплив різних факторів — походження (сировина), технологія виготовлення та експлуатації, які роблять їх артефактами. Детальне дослідження методів та схем розрахунків (в рамках Microsoft Excel) обумовлено, зокрема, вибором лише трьох характерних ознак – питомих активностей гамма-активних нуклідів (ГАН)  $^{40}\text{K}$ ,  $^{228}\text{Ac}$  (ряд  $^{232}\text{Th}$ ),  $^{214}\text{Pb}$  (ряд  $^{238}\text{U}$ ). Необхідність, методологія та приклади такого напрямку досліджень, відомих як пошуковий аналіз, розглянуті в [1, 2]. Використання Excel приведено в [3]; частина матеріалів з цієї теми викладена в [4].

## 1. Методи ідентифікації “зразок-зразок”

Ідентифікація розпочинається з накопичення експериментального матері-

алу, котрий характеризує досліджувані об'єкти (зразки), а потім класифікується за тими чи іншими характерними ознаками.

### 1.1. Апаратурні гамма-спектри

Важливим моментом цього етапу є апріорна визначеність (ім'я) об'єкта (зразка): кожному експериментальному матеріалу ставиться у відповідність ім'я зразка (і, відповідно, сам об'єкт), і навпаки. Так у випадку ПЯГС досліджуваним зразкам ставляться у відповідність їх апаратурні гамма-спектри (АГС), таблиці результатів ідентифікації; в результаті АГС стають ідентифікованими (підписаними). Потім у відповідність ставляться результати кількісного аналізу – таблиці вмістів питомих активностей Ап ГАН та/або концентрації нуклідів, ізотопів, елементів. Таким шляхом створюються атласи (каталоги) АГС.

Більш широка назва цієї процедури створення каталогів – це створення навчальної вибірки.

Наступним кроком ідентифікації є випадки, коли імена зразків невідомі, але є експериментальний матеріал (наприклад їх АГС). Тут виникає обернена проблема встановлення відповідності між АГС і іменем зразка. Якщо є атлас АГС, тоді

задача ідентифікації полягає у послідовному пошуку в ньому АГС, ідентичного АГС, що розглядається. Логіка цих дій зрозуміла – ми можемо оцінити, ідентифікувати зразок тільки методом порівняння з іншими об'єктами, а саме шляхом порівняння між собою характерних ознак об'єктів. Встановлення імен – це, за означенням, робота в номінальних шкалах [5].

## 1.2. Нуклідні спектри зразків

Більш компактно табличний експериментальний матеріал може бути представлений у вигляді нуклідних спектрів зразків (НСЗ) – графічного образу (діаграми) їх питомих активностей ГАН.

НСЗ може бути представлений у вигляді апроксимованого після процедур перевірки ступеня радіоактивної рівноваги, врахування фону вимірів.

Порівняльний аналіз масивів отриманих АГС, таблиць обробки та НСЗ засвідчує, що вони є патернами (в термінології штучного інтелекту) – образами, мітками, маркерами, котрі однозначно ідентифікують зразок.

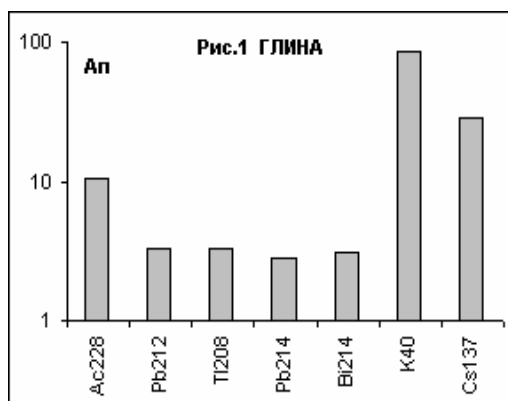


Рис. 1. Нуклідний спектр зразка 2 (глина).  
Ап – питома активності.

На рис. 1 приведено такий нуклідний спектр зразка (тип – глина), який складається з ГАН –  $^{228}\text{Ac}$ ,  $^{212}\text{Pb}$ ,  $^{208}\text{Tl}$  (ряд  $^{232}\text{Th}$ ),  $^{214}\text{Pb}$ ,  $^{214}\text{Bi}$  (ряд  $^{238}\text{U}$ ),  $^{40}\text{K}$ ,  $^{137}\text{Cs}$ .

## 2. Методи групової ідентифікації

### 2.1. Нуклідні спектри типів зразків

Трохи іншою є ситуація, коли виникає проблема групової ідентифікації – ідентифікації не одного об'єкта (зразка), а

множини зразків одного типу серед інших множин (типів). Тут важливо мати дані представлені у вигляді нуклідних спектрів типу зразка (НСТ) – діаграм питомих активностей одного ГАН в множині зразків одного типу.

### 2.2. Стандартизація і нормалізація

Для роботи на рівні групи (типу, множини) необхідно мати свої статистичні шкали, які задаються, зокрема, “нулем” (середніми значеннями) та “одиницею” (відхиленнями). В цих одиницях значення питомих активностей Ап використовуються у вигляді стандартизованих (Аст) і нормалізованих (Ан) значень.

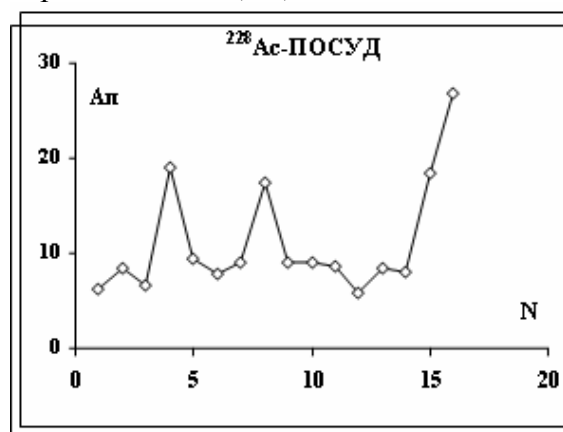


Рис. 2. Нуклідний спектр типу зразка ( $^{228}\text{Ac}$ , посуд) Ап – питома активності. N – номери зразків.

На рис. 2 наведено нуклідний спектр типу зразка Аст  $^{228}\text{Ac}$  (ряд  $^{232}\text{Th}$ ) для зразків посуду.

Об'єднання зразків (елементів) в один тип (множину) загалом є на початку дослідження силовим актом, бо залежить від багатьох чинників – волі дослідника (бажання чи необхідності розглядати множину зразків як один тип), ступені апріорного знання (що досліджувана множина – одного типу). Результатом дослідження має бути факторизація – декомпозиція, чи, навпаки, агрегація, на основі значень факторів, котра відображає істинний стан речей – структурованість масивів значень цих факторів, яка свідчить про структурованість множини зразків.

Маючи каталоги НСТ різних типів зразків, можна виконувати ідентифікацію

невідомих типів (робота в номінальних шкалах).

### 2.3. Варіаційні ряди

Маючи НСТ, можна використати інший тип шкал – ординальні (порядкові) шкали, задавши в НСТ тим чи іншим чином, певне впорядкування. Такі впорядковані НСТ називають варіаційними рядами (ВР).

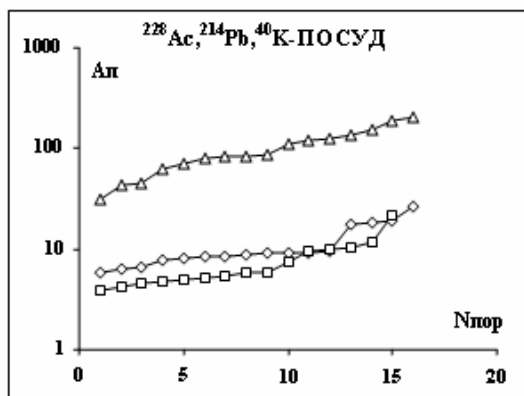


Рис. 3. Варіаційний ряди питомих активностей Ап  $^{40}\text{K}$  (трикутники),  $^{228}\text{Ac}$  (ромби),  $^{214}\text{Pb}$  (квадрати) для зразків посуду. Ап – питома активність (Бк/кг). Nпор – номер впорядкування.

На рис. 3 наведено ВР стандартизованої питомої активності  $^{40}\text{K}$ ,  $^{228}\text{Ac}$  (ряд  $^{232}\text{Th}$ ),  $^{214}\text{Pb}$  (ряд  $^{238}\text{U}$ ) для зразків посуду. Видно, що ВР мають структуру, яка обумовлена індивідуальними та груповими (підтиповими) особливостями – відмінністю різних підтипів.

### 2.4. Розподіли частостей

Маючи таблиці значень питомих активностей, можна визначити ще одну групову характеристику типу: розподіли частостей (РЧ) – відносну кількість (частість) випадків появи того чи іншого значення Ап.

На рис. 4 наведено розподіли частостей Ч питомих активностей Ап, ГАН  $^{40}\text{K}$ ,  $^{228}\text{Ac}$ ,  $^{214}\text{Pb}$  для зразків посуду. Видно, що розподіли частостей відрізняються від нормальних (гаусівських, зокрема) розподілів густин імовірностей.

Аналіз розподілів – це окрема велика тема статистичної обробки даних. Тут ми вкажемо, що аналіз РЧ, виконаний на-

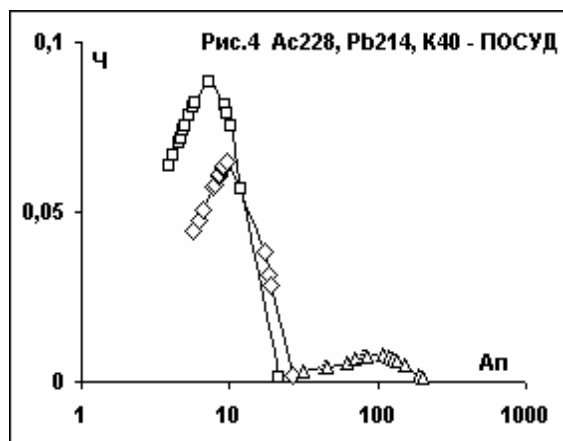


Рис. 4. Розподіли частостей Ч (відносні одиниці) питомих активностей Ап (Бк/кг)  $^{40}\text{K}$  (квадрати),  $^{228}\text{Ac}$  (ромби),  $^{214}\text{Pb}$  (трикутники) для зразків посуду.

ми, підтверджує наявність структури, яка проявляється у ВР.

Порівняльний аналіз масивів отриманих НСТ, ВР, РЧ засвідчує, що вони є патернами типів і можуть бути використані для вирішення певних задач ідентифікації.

### 2.5. Діаграми “Апі-Апj-к”

Точкові та бульбашкові діаграми Excel дозволяють здійснити мапінг [1,2] – дво-, або квазітривимірне представлення даних в графічному вигляді. Типовими представниками мапінгу є власне географічні мапи (карти) та томографії. Діаграми “Апі-Апj-к” питомих активностей Ап ГАНi та ГАНj зразка k (k – ім’я, або, частіше, номер зразка в множині зразків одного типу). За інформативністю діаграми “Апі-Апj-к” не поступаються дендрограмам кластерного аналізу, оскільки дозволяють оцінити дистанцію (відстань) між точками (зразками).

*Діаграми “Апі-Апj/Апl-к”.* Характерними ознаками можуть бути також і відношення питомих активностей  $\text{Апі}/\text{Апl}$ , і, відповідно, діаграми “Апі-Апj/Апl-к”, котрі пов’язують трійку ГАН –  $^{40}\text{K}$ ,  $^{228}\text{Ac}$ ,  $^{214}\text{Pb}$  в різних комбінаціях { i, j, l }.

Порівняльний аналіз масивів отриманих діаграм “Апі-Апj-к”, діаграм “Апі-Апj/Апl-к” та діаграм “Апі/Апj-Апj/Апl-к” засвідчує, що вони також є патернами типів і можуть бути використані для вирішення певних задач ідентифікації.

Діаграми дозволяють виявляти в них як скупчення (кластера) точок, так і тренди (залежності). Це підтверджує факторизацію множини точок – їх декомпозицію (класифікацію) на фактор-множини – сукупності точок, об'єднаних відношеннями подібності (близькості). Крім кластерів, діаграми виявляють фактор-множини точок, об'єднаних іншим типом властивостей – відношеннями функціональної залежності.

Для виявлення та оцінки залежностей між значеннями питомих активностей  $A_{pi} = F(A_{pj})$ ,  $A_{pi} = F(A_{pj} / A_{pl})$ ,  $A_{pi}/A_{pj} = F(A_{pi}/A_{pl})$  використано можливості регресійного аналізу Excel. Дані апроксимувались лінійними залежностями.

## 2.6. Регресійний аналіз

Отримуючи значення коефіцієнта детермінації  $R^2$  (коефіцієнта достовірності апроксимації в термінології Excel), можна виявити такі можливі залежності: шляхом послідовних ітерацій, послідовно виключаючи (або включаючи) із цієї множини значення  $A_{pl}$ , досягають задовільних значень  $R^2$  (напр.,  $R^2 > 0,7$ ), виявляючи в множині точок одну або декілька залежностей.

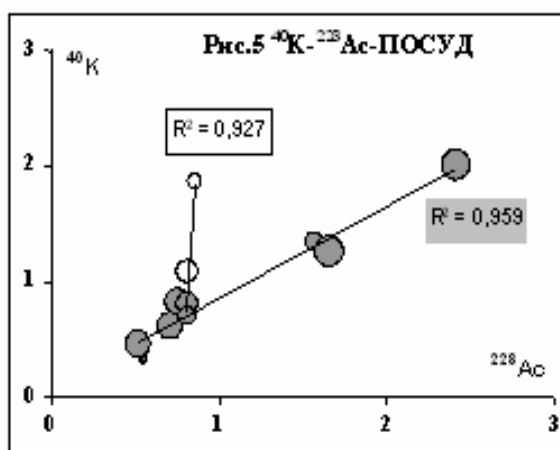


Рис. 5. Діаграма "Апі-Апј-к" ГАН  $^{40}\text{K}$ ,  $^{228}\text{Ac}$  для зразків посуду після регресійного аналізу.  $A_{pl}$  – питомі активності.  $R^2$  – коефіцієнт детермінації. Числа біля значків – номери зразків.

На рис. 5 наведено діаграму "Апі-Апј-к" ГАН  $^{40}\text{K}$ ,  $^{228}\text{Ac}$  для зразків посуду після регресійного аналізу. Чітко прояв-

ляється декомпозиція множини значень  $A_{pl}$  на дві фактор-множини.

## 2.7. Матриці дистанцій

Використовуючи діаграми ми вже фактично використовували не тільки значення питомих активностей  $A_{pl}$ , а і дистанції – відстані між точками: номери (імена) зразків на цих діаграмах. Дистанції  $dA_{pij} = A_{pi} - A_{pj}$  між всіма парами зразків і та ж складають матрицю дистанцій  $dA_{pij}$ .

Значення  $dA_{pij}$  є самостійними величинами, з якими можна працювати в ординальних шкалах, будувати варіаційні ряди  $dA_{pij}$  та розподіли частотей  $dA_{pij}$ .

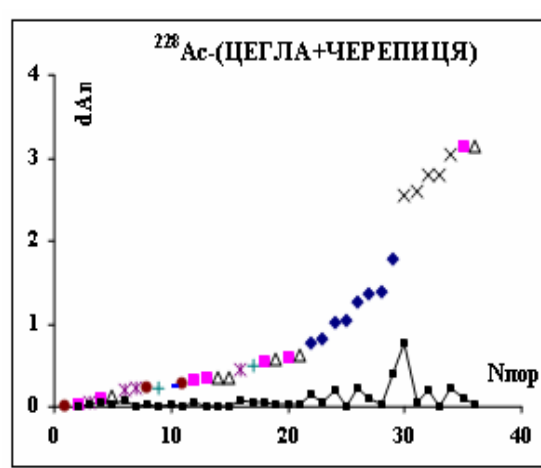


Рис. 6. Варіаційний ряд дистанцій  $dA_{pij}$   $^{228}\text{Ac}$  (верхня крива), та похідна від ряду (нижня крива). Тип зразків – "цегла+черепиця".  $N_{пор}$  – номер впорядкування.

На рис.6 наведено ВР  $dA_{pij}$  для ГАН  $^{228}\text{Ac}$  (тип – "цегла + черепиця") та похідну від цього ряду, котра вказує на наявність структури ВР. Структура ВР – наявність ділянок з різним нахилом – свідчить про структуризацію (факторизацію) дистанцій.

## 2.8. Діаграми матриць дистанцій

Значення дистанцій  $dA_{pij}$  мають графічне зображення на діаграмах "i-j- $dA_{pij}$ ", де i, j – номери (імена) зразків; розміри бульбашок визначаються значеннями  $dA_{pij}$  (чим більші розміри бульбашок, тим більші дистанції  $dA_{pij}$  – відстані між точками).

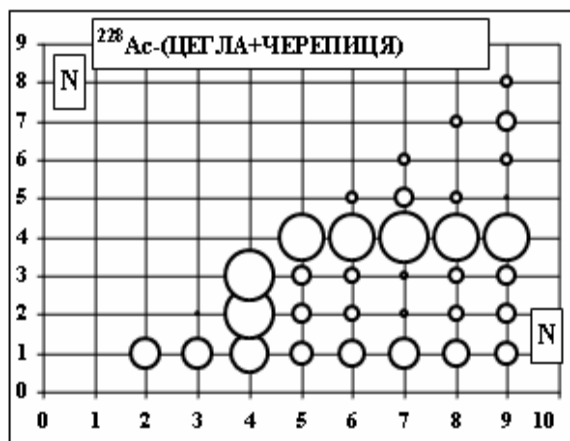


Рис.7. Діаграма “i-j-dAnij”  $^{228}\text{Ac}$ . Тип зразків – “цегла+черепиця”. N – номери зразків i, j.

На рис.7 наведена така діаграма “i-j-dAnij” для ГАН  $^{228}\text{Ac}$  (тип – “цегла+черепиця”). Видно, що діаграми такого типу є достатньо інформативними для вирішення класифікаційних задач.

**2.9. Діаграми дистанцій.** Маючи значення дистанцій dAnij для кожного ГАН, можна побудувати діаграми “dAnij-dAnij-ij” – залежності між значеннями dAnij двох різних ГАН.

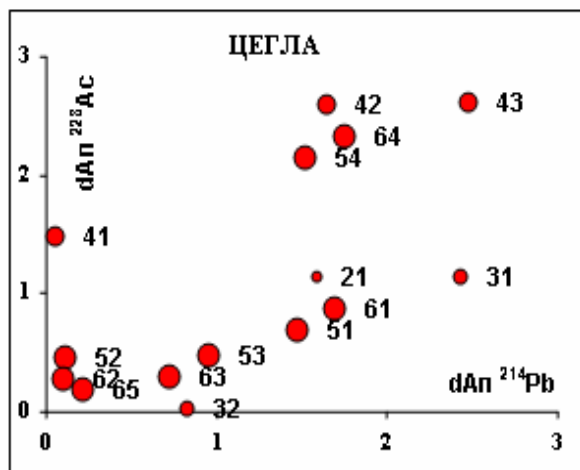


Рис.8. Діаграма “dAnij-dAnij-ij” для  $^{214}\text{Pb}$ ,  $^{228}\text{Ac}$  (тип зразків – цегла). dAnij – дистанції між зразками з номерами i, j (числа біля значків).

На рис.8 наведено діаграму “dAnij-dAnij-ij” для dAnij  $^{228}\text{Ac}$  і dAnij  $^{214}\text{Pb}$  (тип зразків – цегла). Числа біля значків – но-

мери зразків. Видно, як чітко виділяються кластери – скупчення точок.

## Висновки

В роботі нема можливості привести всі отримані нами результати, які демонструють можливості та обмеження розглянутих нами методів. Однак, із приведених нами прикладів, на наш погляд, практична перспективність їх подальшого використання очевидна.

Розглянуті нами методи обробки експериментальних даних не вичерпують всього потужного арсеналу статистичних методів. Однак можна вважати, що використаних нами методів достатньо для виявлення в масивах даних внутрішньої структури (патерни), яка є основою ідентифікації.

Результати різних методів взаємоузгоджені, що свідчить про їх правильність та спільність “походження”, отримані із одних і тих масивів даних.

Очевидною є потреба поступового створення необхідних і зручних через свою наглядність атласів нуклідних спектрів питомих активностей зразків та типів зразків, варіаційних рядів, розподілів частот і т.д.

Вкажемо типові задачі ідентифікації, де можна застосувати досліджені нами методи з використанням ПЯГС  $^{40}\text{K}$ ,  $^{228}\text{Ac}$ ,  $^{214}\text{Pb}$ . Це: розпізнавання “зразок-зразок”; розпізнавання “зразок-тип зразка”; розпізнавання “тип зразка-тип зразка”; розпізнавання “технологія-технологія”; розпізнавання “виріб-сировина (походження)”; розпізнавання “сировина (походження) - сировина (походження)”.

Наші висновки ґрунтуються на відносно невеликому масиві зразків. Зі збільшенням обсягів даних (навчальних виборок) задачі ідентифікації будуть вирішуватись точніше, зокрема така задача, як розпізнавання “технологія-походження”.

### Література

1. А.А. Большаков, Р.Н. Каримов, Методы обработки многомерных данных и временных рядов (Горячая линия-Телеком, Москва, 2007).
2. Ю.Г. Пузаченко, Математические методы в экологических и географических исследованиях (Академия, Москва, 2004).
3. А.Н. Васильев, Научные вычисления в Microsoft Excel (Диалектика, Москва, 2004).
4. М.В. Стец, В.Т. Маслюк, В.М. Бузаш, О.М. Матевка. Тез. докладов 7 конференции по физике высоких энергий, ядерной физике и ускорителям (Харьков, 2009), с.32.
5. Н.В.Хованов. Математические основы теории шкал измерения качества (Изд-во ЛГУ, Ленинград,1982).

## APPLIED NUCLEAR GAMMA-SPECTROMETRY IN CERTAIN PROBLEMS OF IDENTIFICATION OF CERAMIC PRODUCTS

**O.M. Matyovka<sup>1</sup>, M.V. Stets<sup>2</sup>, V.T. Maslyuk<sup>2</sup>, V.M. Buzash**

<sup>1</sup> Uzhgorod National University, Uzhgorod

<sup>2</sup> Institute of Electron Physics, Ukr. Nat. Acad. Sci., Uzhgorod

Based on the specific activities An of  $^{40}\text{K}$ ,  $^{228}\text{Ac}$   $^{232}\text{(Th series)}$ ,  $^{214}\text{Pb}$  ( $^{238}\text{U series}$ ) in environmental samples (i.e. soils, clays and relevant ceramic products) the possibility of their use in different statistical processing methods (frequency dependence analysis, variation series, Anj-k, An/An-An/An diagrams, distance matrices) has been estimated in order to solve the identification problems. The consistency of the results of the above methods confirms the correctness of factorization of a variety of An quantities for different types of the samples.

