

Intelligent Multisensor System for Analytical Control of Sausages

A.A. Kalinichenko*, L.U. Arseniyeva

National University of Food Technologies, 68 Volodymyrska str., Kyiv, Ukraine, 01601;

*e-mail: asya.kalini4enko@gmail.com

Received: January 19, 2019; Accepted: July 02, 2019

DOI: 10.17721/moca.2019.57-72

The new technique of intelligent analysis of chemical aroma patterns of boiled sausages obtained by the electronic nose for authentication and microbiological safety assessment is developed. The informativeness of features extracted from steady-state responses of the multisensor system and robustness of chemometric algorithms for solving the objectives of qualitative and quantitative analysis of sausage volatile compounds are investigated. The classification model was built using maximum response values as input vectors of an optimized probabilistic neural network, which allows obtaining a 100 % accuracy of different sample grades identification and detection samples adulterated with soy protein. The method of partial least squares regression and area values as features were used for regression modelling and prediction of QMAFAnM with a relative error less than 12 % for a microbiological safety assessment of previously identified sausages. The use of the robust analytical technique to assess authentication, adulteration, total bacterial count for one measurement using the electronic nose in combination with machine learning algorithms will allow to significantly reduce the measurement time and the cost of analysis, and avoid subjective estimation of the results.

Keywords: electronic nose, chemical pattern recognition, classification, probabilistic neural network, prediction, partial least squares regression

Інтелектуальна мультисенсорна система для аналітичного контролю ковбасних виробів

А.О. Калініченко*, Л.Ю. Арсеньєва

Національний університет харчових технологій, вул. Володимирська 68, Київ, Україна, 01601;

*e-mail: asya.kalini4enko@gmail.com

Надійшла: 19 січня 2019 р; Прийнята: 02 липня 2019 р

DOI: 10.17721/moca.2019.57-72

Розроблено методику інтелектуального аналізу хімічних образів запахів електронного носа для оцінки автентичності та мікробіологічної безпечності варених ковбасних виробів. Досліджено інформативність параметрів, вилучених з статичних відгуків мультисенсорної системи та робастність хемометричних алгоритмів для вирішення задач якісного та кількісного аналізу складу летких сполук ковбасних виробів. Розроблено класифікаційну модель з використанням максимальних сигналів сенсорів як вхідних векторів оптимізованої імовірнісної нейронної мережі, що дозволяє з 100 % точністю ідентифікувати зразки різного сорту та виявити фальсифіковані зразки соєвим ізолятом. Розроблені регресійні моделі з використанням площ кривих відгуків сенсорів та методу проєкції на латентні структури для оцінки мікробіологічної безпечності ідентифікованих зразків з відносною похибкою прогнозування показника КМАФАнМ менше 12 %. Використання робастної методики оцінки автентичності, фальсифікації, загального мікробного показника ковбасних виробів за одне вимірювання з використанням електронного носа в поєднанні з машинним навчанням дозволить суттєво зменшити тривалість та знизити собівартість аналізу, уникнути суб'єктивності оцінки результатів.

Ключові слова: електронний ніс, розпізнавання хімічних образів, класифікація, імовірнісна нейронна мережа, прогнозування, регресія на латентні структури

Методологія аналізу багатокомпонентних газових сумішей, в тому числі харчових продуктів, з використанням електронного носа дозволяє вилучити великий об'єм даних, що містять

інформацію про властивості об'єкта та особливості вимірювальної системи. Селективність аналізу мультисенсорних систем з перехресною чутливістю сенсорів забезпечується багатоканальністю

надходження інформації. Низька селективність п'єзосенсорів не становить проблеми для ідентифікації та навіть кількісного аналізу за наявності хоча б декількох сенсорів в масиві з високою масовою чутливістю до певних груп або класів сполук. Багатовимірний відгук системи представляє собою характерний відбиток, або хімічний образ запаху зразка. Під хімічним образом мається на увазі математичний портрет будь-якого стану сенсорної системи в момент або на кінець експозиції молекул-аналітів.

Основною задачею аналізу багатовимірних відгуків електронного носа є пошук та вилучення інформації в базах даних, що зазвичай здійснюється з використанням методів обробки багатовимірних даних (data mining).

Проблема застосування хеометричних методів полягає в тому, що вирішення конкретної практичної задачі потребує нового пошуку та вилучення необхідних ознак та параметрів, моделювання, вибору оптимального хеометричного методу обробки багатовимірних даних певної структури та оптимізації алгоритмів [1–5].

На сьогодні відсутня методологічна схема інтелектуального аналізу даних п'єзоелектронного носа для одночасної оцінки якісних та кількісних показників.

Актуальним є розроблення методологічної схеми аналізу хімічних образів запахів п'єзоелектронного носа та надання рекомендацій з вилучення інформативних параметрів, синтезу та моделювання алгоритмів машинного навчання для вирішення різних задач якісного та кількісного аналізу харчових продуктів. Отримані моделі повинні бути підкріплені надійною перевіркою з розумінням меж застосування моделей для ефективного практичного використання.

Харчовий продукт є складним аналітичним об'єктом за рахунок мінливості фізико-хімічного стану системи, що потребують комплексного аналізу показників. Склад летких сполук системи залежить від технологічних параметрів, якості вихідної сировини, а також суттєво змінюється в процесі зберігання. Розроблення ефективних для практичного використання методик аналітичного контролю сукупності параметрів потребує надання рекомендацій з аналізу хімічного образу запаху зразка для оцінки різних характеристик об'єкта за одне вимірювання, оптимізації алгоритмів та розроблення моделей, що будуть робастними, тобто нечутливими до відхилень та неоднорідностей у вибірці, можливих викидів. Тому для вирішення задачі ідентифікації ковбас використовували зразки різної якості, а під час аналізу безпечності досліджували зразки різного складу для отримання робастних моделей.

Актуальність розроблення експрес-методик оцінки автентичності, якості та безпечності, виявлення фальсифікації ковбасних виробів

з використанням мультисенсорних систем аргументується в роботах авторів [6–8], а також в роботах світових дослідників [2, 9, 10].

Відповідно до стандарту ДСТУ 4436 варені ковбаси вищого сорту не повинні містити білкових препаратів (соевих, молочних), в рецептурі ковбас нижчих сортів (1-3 сортів) дозволено використовувати соєві компоненти у кількостях 10-20 мас.% [11]. Однак, виробники, замінюючи м'ясну сировину соєвими продуктами, зазвичай не вказує масовий вміст замінильника та ковбасні вироби зазнають фальсифікації. На сьогодні відсутня методика експрес-виявлення та визначення соєвих компонентів в варених ковбасних виробах [12–16].

Для оцінки мікробіологічної безпечності ковбас необхідно понад 72 год., що робить неможливим скринінг об'єктів. Визначення показника загального мікробного числа, або кількості мезофільних аеробних та факультативно-анаеробних мікроорганізмів (КМАФАнМ, КУО/г) ковбасних виробів за арбітражною методикою є трудомістким та суб'єктивним аналізом, що потребує багатоетапної підготовки проби та поживного середовища [17].

Активно використовують хімічні сенсорні системи для встановлення ідентичності та відмінностей у складі газової фази, кількісної оцінки вмісту груп або класів летких сполук, в т. ч. продукуючих мікроорганізмами під час псування харчових продуктів [2, 10, 18–20]. Однак, методики є специфічні та направлені на вирішення однієї вузької задачі.

Таким чином, метою роботи є розроблення робастної методики (як на хімічному, так і на хеометричному рівні) оцінки автентичності ковбасних виробів, виявлення фальсифікації соєвим ізолятом та прогнозування безпечності (показника КМАФАнМ) за одне вимірювання з використанням інтелектуального електронного носа, що включає:

- дослідження інформативності похідних параметрів з відгуків сенсорів в задачах якісного та кількісного аналізу складу летких сполук ковбас;
- побудову класифікаційних та регресійних моделей, та оптимізацію алгоритмів машинного навчання;
- розроблення методологічної схеми інтелектуального аналізу хімічних образів п'єзоелектронного носа для оцінки якісних та кількісних показників ковбасних виробів за одне вимірювання.

Експериментальна частина

Готували стандартні зразки ковбасних виробів за рецептурою вареної ковбаси Лікарської вищого сорту. Зразки нижчих сортів та фальсифіковані проби готували шляхом заміни м'ясної суміші (70 % свинини та 25 % яловичини) на ізолят соєвого білка в масових концентраціях 10, 20, 30 % (в гідратованій формі). Ізолят соєвого білка гідратували дистильованою водою у співвідношенні 1:6. Ізолят соєвого білка (білок $\geq 90.0\%$, вологість

$\leq 8.0\%$, $pH = 7.0 \pm 0.5$) був придбаний в компанії «Протеїн. Технології. Інгредієнти» (Митищі, РФ). Зразки зберігали за температури $20 \pm 1^\circ\text{C}$ протягом 4 діб. Отримані зразки заздалегідь відомого складу використовували як еталонні зразки для формування бази даних хімічних образів з метою ідентифікації ковбас. Вимірювали та контролювали відповідність фізико-хімічних та мікробіологічних показників ковбасних виробів вимогам стандартів.

Визначали показник загального мікробного числа (КМАФАнМ, КУО/г) протягом зберігання зразків для оцінки безпечності ковбас, використовуючи арбітражну методику [17], для подальшої побудови калібрувальних моделей.

Загалом було аналізовано 48 зразків ковбасних виробів у трьох повторностях, масив даних з 144 вимірювань використовували для дослідження варіації всередині реплік під час обробки хемометричними методами.

Електронний ніс (ТОВ «Сенсоріка-Нові Технології», Вороніж, РФ) використовували для отримання хімічних образів запахів ковбасних виробів. Виготовляли сенсори на базі кварцових резонаторів АТ-зрізу з основною частотою коливання 10 МГц з срібними електродами діаметром 5 мм (БАТ «П'єзо», Москва, РФ). Для створення покриттів сенсорів використовували розчини хроматографічних фаз в ацетоні та хлороформі, що характеризуються різною полярністю, стабільністю та селективністю по відношенню до летких маркерів м'ясних та соєвих продуктів: поліетиленгліколь себацінат (PEG seb – сенсор 1), поліетиленгліколь адипінат (PEG ad – сенсор 2), дициклогексано-18-краун-6 (DCG18k6 – сенсор 3), Triton X-100 (TX-100 – сенсор 4), поліетиленгліколь 2000 (PEG-2000 – сенсор 5), полідіетиленгліколь сукцинат (PDEG suc – сенсор 6), полівінілпіролідон (PVP – сенсор 7), виготовлені компанією Sigma-Aldrich Chemie GmbH (ТОВ «Алсі», Київ, Україна). Формували покриття сенсорів масою 10-20 мкг способом статичного випаровування каплі.

Експериментальні умови статичного парово-фазного аналізу ковбас з використанням електронного носа: проби масою 5 ± 0.1 г поміщували в пробовідбірники об'ємом 60 см^3 , індивідуальним шприцем відбирали 3 см^3 рівноважної газової фази (РГФ) та вводили в комірку детектування зі швидкістю $60\text{ см}^3\text{хв}^{-1}$. Час вимірювання складав 60 с з інтервалом 1 с. Комірку детектування електронного носа очищували потоком осушеного повітря протягом 3 хв. до інжекції нової проби.

З вихідних статичних відгуків сенсорів вилучали похідні параметри та досліджували їх інформативність для вирішення задач якісного та кількісного аналізу, а саме:

- площі кривих одиничних сенсорів та осі абсциси часу їх обмежуючих (S_i , Гц $\times$ с);

- максимальні зміни частот коливання сенсорів за час вимірювання ($\Delta F_i^{\max}$, Гц).

Параметр $\Delta F_i^{\max}$ представляє максимальну ступінь реакції сенсора на пробу газової фази, в той час як S_i характеризує всю криву відгуку та відображає кінетичні аспекти адсорбції летких сполук на покритті сенсора (взаємодію летких сполук з плівкою сенсора та дифузію на поверхні). Відмінності у формі кривих відгуків містять інформацію про чутливість та селективність сенсорів до летких маркерів або груп летких сполук.

Формували багатовимірні масиви (144 вимірювання $\times$ 7 параметрів сенсорів), що поділені на 4 класи (групи) залежно від вмісту соєвого ізоляту (CI): 1 клас – зразки вареної ковбаси в. с. (0% CI), 2 клас – зразки з 10 мас. % CI, 3 клас – зразки із 20 мас. % CI, 4 клас – зразки із 30 % CI для подальшого аналізу методами обробки багатовимірних даних.

Класифікаційні моделі будували з використанням імовірнісної нейронної мережі, реалізованої в програмному пакеті MATLAB R2014b (MathWorks Inc., Natick, USA). Для візуалізації та аналізу структури даних, кластеризації та стиснення даних застосовували метод головних компонент, регресійні моделі будували із використанням методу проєкції на латентні структури, обидва алгоритми реалізовані в програмному забезпеченні The Unscrambler 10.4 software (CAMO Software AS., Oslo, Norway).

Результати та обговорення

Моделювання класифікаційної моделі для оцінки автентичності ковбасних виробів та виявлення фальсифікації соєвим ізолятом

У разі аналізу окремих летких сполук або двох-, трьохкомпонентних сумішей хімічний образ ототожнюється з запахом цих сполук, хімічним складом, а також з хімічного образу можна вилучити інформацію про вміст кожного компонента. Однак, під час аналізу багатоконпонентних сумішей під хімічним образом слід розуміти фізико-хімічну структуру проби (взаємодія її компонент, розташування частинок в просторі, тощо), візуальний профіль багатоконпонентної суміші не є сумою профілів окремих сполук в певних концентраціях.

Отже, тренування електронного носа для ідентифікації харчових продуктів потрібно проводити на еталонних зразках відомого складу для формування бази даних хімічних образів. Формування проб та їх компонентний склад детально наведено в п.2. Аналізували зразки ковбасних виробів без CI (ковбаси вищого сорту), зразки з 10 мас. % CI (ковбаси першого сорту), із 20 (ковбаси третього сорту) та 30 мас. % замітника (фальсифіковані зразки).

Класифікаційну задачу логічно розпочати з візуалізації експериментальних даних та аналізу

структури дисперсії методом головних компонент (рис. 1). МГК є найбільш популярним методом кластеризації даних, а також основою для побудови багатьох методів класифікації, таких як метод формального незалежного моделювання аналогій класів SIMCA (Soft Independent Modeling of Class Analogy), лінійний та квадратичний дискримінантний аналіз (Linear and Quadratic Discriminant Analysis, LDA, QDA), дискримінантний аналіз за допомогою регресії на латентні структури (Partial Least Squares – Discriminant Analysis, PLS-DA) [21].

МГК одночасно вирішує дві важливі задачі: переведення даних в більш придатну систему координат (початок якої лежить в центрі області даних) та зменшення розмірності (використання тільки необхідної кількості головних компонент (ГК), які відображають структуру даних). Число головних компонент, необхідних для моделювання даних, вибирають за відсотком поясненої ними дисперсії, однак не завжди потрібно прагнути 100%, оскільки в такому випадку можна включити шумову складову.

В даній роботі матрицю даних з сімома параметрами стиснули до набору узагальнених координат x , y , z та трьох головних компонент. ГК1, ГК2 та ГК3 пояснюють 97% дисперсії під час декомпозиції матриці параметрів $\Delta F_i^{\max}$ та 96% у випадку матриці параметрів S_i .

За результатами оброблення матриці параметрів $\Delta F^{\max}$ (рис. 1) можна виділити три кластери: перший – зразки вареної ковбаси в. с. без додавання заміни, другий – зразки з 10% мас. СІ, третій кластер об'єднує зразки з 20 та 30% мас. заміни м'ясної сировини на соєвий ізолят.

На графіку рахунків матриці параметрів S_i можна виділити два кластери даних: перший – зразки вищого сорту (0% СІ) та ковбаси 10% мас. СІ, другий кластер об'єднує зразки з 20 та 30% мас. СІ. Всі типи ковбасних виробів дискримінувати з використанням МГК неможливо. Отримані результати за відсутності викидів та високого відсотка пояснювальної дисперсії вказують на нелінійну структуру даних.

Неефективним є використання лінійних методів класифікації, основою яких є МГК, якщо дані є суттєво нелінійними. В таких випадках доцільно використовувати нелінійні функції та методи оброблення даних, такі як нейронні мережі. В даній роботі матрицю рахунків МГК використовували в якості вхідних векторів нейронної мережі. В такому випадку, попереднє оброблення даних методом головних компонент дозволить отримати швидку в навчанні нейронну мережу зі зменшеною кількістю вхідних нейронів.

Загалом попередня обробка даних спрямована на підвищення ефективності функціонування нейронної мережі за рахунок приведення даних в необхідний формат. Слід зазначити, що використання необхідної попередньої обробки даних залежить як від виду та структури даних, так і від архітектури нейронної мережі. Одним із завдань було дослідити необхідність попередньої обробки даних електронного носа для найкращої їх апроксимації із використанням імовірнісної нейронної мережі.

Отже, в роботі використовували три види оброблення та трансформування даних: нормалізацію, автошкалування та оброблення методом головних компонент (МГК).

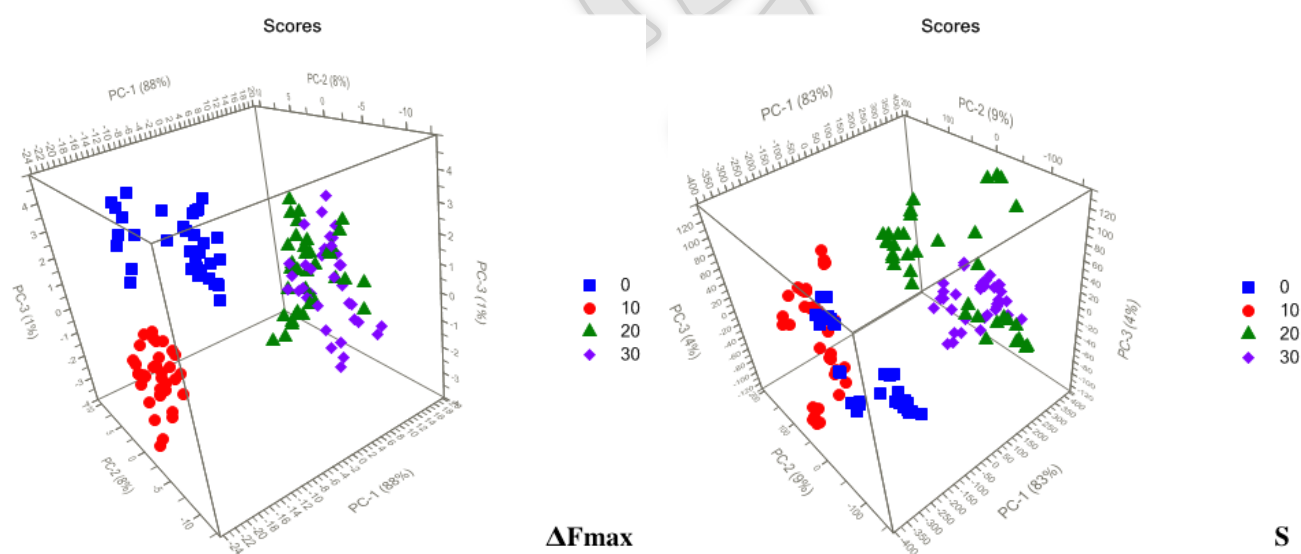


Рис. 1. 3D графіки рахунків МГК під час аналізу ковбасних виробів із різним вмістом соєвого ізоляту за вихідними параметрами електронного носа.

Імовірнісна нейронна мережа (Probabilistic neural networks, PNN) є надійним та швидким алгоритмом класифікації з навчанням. PNN характеризується простою архітектурою та складається з чотирьох шарів: вхідного, двох проміжних шарів (шару зразків та підсумовуючого) та вихідного шару нейронів. В програмному пакеті Matlab проміжні шари називаються радіально-базисний та конкуруючий відповідно [22]. Опис архітектури імовірнісної нейронної мережі для класифікації чотирьох категорій зразків за вихідними параметрами електронного носа та описання типу функції активації мережі наведено в роботі [6].

PNN дозволяє виразити нелінійні комбінаційні ефекти всіх вхідних параметрів як єдину функцію, яка деталізує зміни складу харчового зразка. Це досягається оптимізацією значення відхилення функції Гауса, що визначає ширину "дзвону" з центром в кожному кластері та ступінь інтерполяції між точками.

Навчання будь-якої нейронної мережі починається з формування навчальної та тестової вибірок. Досліджуваний масив даних рандомним чином поділяли на навчальну та тестову вибірки різного співвідношення. Під оптимальним об'ємом навчальної вибірки розуміли таке число зразків, яке забезпечувало максимальну надійність класифікації зразків тестової вибірки. Коефіцієнт L , % показує, яка частка зразків від їх загального числа знаходиться в навчальній вибірці:

$$L = \frac{H}{M} \times 100\%, \quad (1)$$

де H – число зразків навчальної вибірки, M – загальне число зразків.

В роботі ненадійність класифікації оцінювали як частку невірно класифікованих зразків тестової вибірки [23]:

$$U = \frac{t}{T} \times 100\%, \quad (2)$$

де t – число невірно класифікованих зразків тестової вибірки, T – загальне число зразків тестової вибірки.

Оптимальний та представницький об'єм навчальної вибірки для вихідних параметрів електронного носа (S_i та $\Delta F_i^{\max}$) складав 75 %, тестової вибірки – 25 %.

Надійність класифікації оцінювали як:

$$R = 100 - U, \quad (3)$$

Отримані класифікаційні моделі перевіряли за допомогою техніки 4-кратної перехресної перевірки, а під надійністю класифікації розуміється середнє

значення точності класифікації 4 циклів навчання та тестування.

Параметрична оптимізація імовірнісної нейронної мережі полягає в визначенні параметра згладжування σ , що контролює коефіцієнт масштабування експоненціальної функції активації.

Експериментальним шляхом не складно визначити оптимальне значення σ , так як швидкість неправильної класифікації змінюється поступово за невеликих змін σ . Рекомендують обирати значення ступеня згладжування як функцію розмірності вимірювального простору та кількості зразків навчальної вибірки [24]. За налаштуванням в алгоритмі PNN Matlab $\sigma = 0.1$.

Встановлено, що найвища надійність дискримінації тестових зразків варених ковбас (95.8 %) за вихідним параметром електронного носа S_i без попереднього оброблення може бути отримана за будь-яких значень σ в достатньо широкому інтервалі від 3.6 до 54.0. Крім того, спостерігаємо незначне погіршення точності класифікації зразків за значень σ в інтервалах від 2.4 до 3.5 та від 54.1 до 100.0.

З наведених результатів в табл. 1 можна зробити висновок, що використання процедур попередньої обробки даних не завжди є виправданим, оскільки зашумлені змінні можуть здійснювати не виправдано великий вплив, або якщо вибіркові значення дисперсії більш-менш однакові для всього набору змінних, в такому випадку відсутня необхідність попередньої обробки. Нормалізація за мінімальним та максимальним сигналами змінює вагу вкладу різних змінних. Наприклад, значення параметра площі під кривою (S_i) найбільш полярного сенсора з покриттям PVP під час нормалізації за рядками приймається як максимальне значення інтервалу 1 та відповідно його суттєвий вплив на дискримінацію зразків нівелюється, за рахунок чого точність класифікації знижується.

Метод головних компонент передбачає лінійну структуру даних. Застосування алгоритму пов'язано з втратою дисперсії даних внаслідок їх лінеаризації, що стає причиною обмеженого числа факторних ознак, що включаються в модель. Однак, попереднє оброблення даних МГК дозволяє знизити їх розмірність з семи векторів до трьох (ПД=96–97 %) та збільшити швидкість навчання нейронної мережі.

Максимальну надійність (100 %) результатів багатокласової класифікації зразків варених ковбас за параметром $\Delta F_i^{\max}$ отримали в інтервалі значень σ від 0.2 до 1.0. Використання попереднього оброблення та трансформації даних незначно погіршило точність ідентифікації зразків.

Таблиця 1. Результати прогнозування 4 класів ковбас з використанням PNN та параметрів електронного носія.

Попередня підготовка даних	Відхилення функції (σ)	Ненадійність класифікації (U , %)	Надійність класифікації (R , %)
S_i			
Вихідні дані	3.6–54.0	4.2	95.8
Нормалізація	0.1–0.2; 0.9	8.3	91.7
Масштабування	0.1	6.2	93.8
МГК (центрування)	3.0–3.1	6.3	93.7
	3.4–3.8	6.2	93.8
$\Delta F_i^{\max}$			
Вихідні дані	0.2–1.0	0	100
Нормалізація	0.1–1.0	4.2	95.8
Масштабування	0.7–0.8	2.1	97.9
МГК (центрування)	0.2–1.0	4.2	95.8

Моделювання регресійних моделей для прогнозування загального мікробного числа ковбасних виробів

Отже, використання матриці параметрів $\Delta F_i^{\max}$ як вхідних векторів імовірнісної нейронної мережі дозволило надійно класифікувати чотири категорії ковбасних виробів, оцінити автентичність ковбасних виробів та виявити фальсифіковані зразки для подальшого аналізу їх безпечності.

Досліджували динаміку зміни мікробіологічного показника КМАФАнМ (загальної кількості мікроорганізмів) під час зберігання ковбасних виробів з використанням арбітражного методу [17]. На інтенсивність перебігу мікробіологічних процесів в ковбасних виробках впливають склад та властивості сировини, початкова контамінація ковбас мікроорганізмами та температура зберігання зразків. Відповідно до стандарту, кількість МАФАнМ в варених ковбасних виробках вищого та першого сортів допускається не більше 1.0×10^3 КУО/г, в ковбасах третього сорту – 5.0×10^3 КУО/г [11]. Результати, наведені в табл. 2, свідчать, що додавання соєвих компонентів пришвидшує мікробне псування виробів, що пов'язано з особливостями мікрофлори соєвих продуктів.

Отже, будували окремі багатовимірні регресійні моделі прогнозування загального мікробного числа ковбасних виробів без додавання соєвих компонентів та із вмістом СІ в межах норм стандарту (10–20 % мас.), що пов'язано з різною мікрофлорою м'ясних та соєвих продуктів. Псування м'ясо-соєвих виробів характеризується більшою швидкістю зростання мікроорганізмів, що цілком ймовірно пов'язано з синергетичною взаємодією різних штамів мікроорганізмів в ковбасному виробі [25]. Специфічні групи мікроорганізмів соєвих

продуктів продукують більшу кількість метаболітів, накопичення яких також сприяє пришвидшенню псування як соєвих, так і м'ясних складових ковбасних виробів [25–27].

Найбільш популярним методом багатовимірної калібрування в хемометриці є метод проєкції на латентні структури (partial least squares regression, PLSR), який здійснює одночасну декомпозицію матриць X (матриці предикторів, параметрів сенсорів) та Y (матриці відгуків, значень прогнозованого мікробіологічного показника, КМАФАнМ). Проєкції будуються таким чином, щоб максимізувати кореляцію X та Y рахунків, що дає змогу описувати складні зв'язки, використовуючи меншу кількість головних компонент [21]. Однією з умов успішного використання PLS є лінійність даних. Однак, Y може бути лінійною функцією від комбінації декількох нелінійних змінних X , тому нелінійна структура даних матриці предикторів може не стати проблемою під час побудови багатовимірної регресійної моделі.

Значення мікробіологічного показника розподілені асиметрично, тому застосовували логарифмічне перетворення вихідних даних для подальшого аналізу. Таким чином, під час побудови моделей використовували натуральний логарифм загального мікробного числа ковбас ($\ln$ КМАФАнМ).

Для першому етапі побудови регресійної моделі для прогнозування загального мікробного числа ковбасних виробів в.с. (без додавання СІ) формували матрицю предикторів, які б суттєво корелювали з мікробіологічним показником, вилучали предиктори із значеннями коефіцієнтів кореляції нижче критичного значення коефіцієнта кореляції Пірсона для набору даних $r_{36, 0.05} = 0.33$ (табл. 3).

Таблиця 2. Динаміка зміни мікробіологічного показника ковбасних виробів.

Зразки ковбаси	Тривалість зберігання зразків, год. × 12						
	1	2	3	4	5	6	7
КМАФАнМ, КУО/г							
Вищий сорт 0 мас.% CI	$2.3 \cdot 10^2$	$3.2 \cdot 10^2$	$1.0 \cdot 10^3$	$1.1 \cdot 10^3$	$2.4 \cdot 10^3$	$6.3 \cdot 10^3$	$9.9 \cdot 10^3$
10 мас.% CI	$2.7 \cdot 10^2$	$3.2 \cdot 10^2$	$1.0 \cdot 10^3$	$1.5 \cdot 10^3$	$1.3 \cdot 10^4$	$1.4 \cdot 10^4$	$2.2 \cdot 10^4$
20 мас.% CI	$5.0 \cdot 10^2$	$5.7 \cdot 10^2$	$1.5 \cdot 10^3$	$2.0 \cdot 10^3$	$1.4 \cdot 10^4$	$1.6 \cdot 10^4$	$2.8 \cdot 10^4$
30 мас.% CI	$1.5 \cdot 10^2$	$3.0 \cdot 10^2$	$1.5 \cdot 10^3$	$2.0 \cdot 10^3$	$1.6 \cdot 10^4$	$2.1 \cdot 10^4$	$4.8 \cdot 10^4$

Таблиця 3. Значення коефіцієнтів кореляції параметрів електронного носа з мікробіологічним показником КМАФАнМ ковбас в.с.

Параметри відгуків сенсорів	r	Параметри відгуків сенсорів	r
$\Delta F^{\max}$ PEGseb	-0.73	S PEGseb	-0.65
$\Delta F^{\max}$ PEGad	-0.67	S PEGad	-0.53
$\Delta F^{\max}$ DCG18k6	0.76	S DCG18k6	0.94
$\Delta F^{\max}$ TX-100	0.87	S TX-100	0.79
$\Delta F^{\max}$ PEG-2000	0.23	S PEG-2000	-0.67
$\Delta F^{\max}$ PDEGsuc	0.67	S PDEGsuc	-0.89
$\Delta F^{\max}$ PVP	-0.92	S PVP	-0.92

Таким чином, з матриці предикторів вилучили неінформативний параметр $\Delta F^{\max}$ PEG-2000. Найбільш сильно корелює з показником ln КМАФАнМ параметри сенсорів з покриттями полівінілпіролідон, полідиетиленгліколь сукцинат та дициклогексано-18-краун-6, що проявляють високу чутливість до парів амінів та кислот [1, 28], що продукуються мікроорганізмами під час псування ковбас. Потрібно зазначити, що одновимірний кореляційний аналіз є лише розвідувальним дослідженням, та результати можуть відрізнятися від взаємозв'язку X та Y змінних в просторі рахунків багатовимірної калібрувальної моделі.

Моделювали три регресійних моделі для порівняння інформативності X-матриць предикторів для прогнозування загального мікробного числа (Y):

- PLS1 з використанням семи параметрів S_i та шести $\Delta F_i^{\max}$;
- PLS2 з сімома параметрами S_i електронного носа;
- PLS3 з використанням шести параметрів $\Delta F_i^{\max}$.

В багатовимірному регресійному аналізі завжди використовується надмірна кількість параметрів, що часто стає причиною побудови нестабільних моделей, непридатних для практичного застосування.

Важливо запропонувати оптимальну кількість предикторів, що містять інформацію про кількість продукуючих метаболітів, що в свою чергу має бути безпосередньо пов'язано із збільшенням загальної

кількості мікроорганізмів протягом псування ковбас.

Матриці даних X та Y автошкалували ($1/S_{dev}$) для отримання одиної дисперсії всіх змінних та уникнення домінуючого впливу на модель через порядок величини.

Об'єм навчальної вибірки для побудови PLSR моделей прогнозування ln КМАФАнМ ковбасних виробів в.с. (0% мас. CI) становив 24 зразка. Всі калібрувальні моделі перевіряли за допомогою перехресної перевірки з числом сегментів 12.

За результатами аналізу пояснювальної дисперсії обрано кількість значущих головних компонент (ГК) або факторів для кожної моделі: PLS1 – 1 ГК (62%, 95%), PLS2 – 1 ГК (65%, 95%), PLS3 – 3 ГК (89%, 94%), де в дужках зазначені відсоток X та Y дисперсій. Перші компоненти описують найбільші зміни значень в Y, тому є найбільш інформативними. Зазначимо, що додавання надлишкових компонент покращує точність калібрувальних моделей PLS1 та PLS2, однак ГК2 описує тільки 1% варіації Y в обох випадках. Для майбутнього прогнозування більш важливим є отримання робастних моделей, що описують системні зміни в даних. Проблема збалансованого опису даних також розглядається в роботах А. Хоскюдссона, який запропонував концепцію моделювання – Н-принцип, згідно якого поліпшення точності моделювання спричиняє погіршення точності прогнозування [29].

Для обраної кількості ГК розраховуються значення регресійних коефіцієнтів ($\hat{b}_0$ та $\hat{b}$), що будуть використовуватися в рівнянні прогнозування показника Іп КМАФАНМ:

$$\hat{y} = 1\hat{b}_0 + X\hat{b}. \quad (4)$$

Регресійні коефіцієнти підсумовують взаємозв'язки між усіма предикторами X та відкликом Y , надають інформацію про вагомість вкладу кожної X -змінної в регресійну модель (табл. 4). У випадку PLS3 підсумовуються взаємозв'язки між предикторами та відкликом для трьох факторів разом. Зважені регресійні коефіцієнти (b) враховують ваги, що застосовувалися до предикторів, в даному випадку $1/SD_{dev}$ (автошкалювання).

Найбільш важомі коефіцієнти для всіх моделей є параметри сенсорів з покриттями дициклогексано-18-краун-6 та тритон X-100 (позитивна кореляція з показником Іп КМАФАНМ), а також полівінілпіролідон та полідиетиленгліколь сукцинат (від'ємна кореляція з показником Іп КМАФАНМ). Зважені регресійні коефіцієнти PLS1 мають майже однаковий вплив на модель, PLS2 та PLS3 містять домінуючі предиктори, однак перерахунок моделей без менш значущих параметрів непотрібен.

Точність калібрування та перехресної перевірки прийнято характеризувати значеннями середньоквадратичних помилок (RMSE cal та RMSE val). Відповідно, чим менше RMSE cal та

RMSE val, тим точніше описуються дані. Крім того, якість калібрування та перехресної перевірки характеризується квадратами коефіцієнтів кореляції R^2 cal та R^2 val, які розраховуються в масштабі від нуля до одиниці, де нуль вказує на відсутність кореляції, а одиниця відображає відмінну лінійну кореляцію.

На практиці регресійні моделі часто переоцінюють, точність опису калібрувальних даних буває значно краще, ніж точність прогнозування. Таким чином, точність моделювання (RMSE cal) та прогнозування (RMSEP) потрібно завжди розглядати разом.

Отримані прогностичні моделі оцінювали за тестовим набором з 12 нових зразків. RMSEP є середньоквадратичною помилкою прогнозування тестових зразків (5). Значення Іп КМАФАНМ в процесі зберігання ковбас в.с. змінювалися в межах від 5.44 до 9.20, таким чином відносна похибка RMSEP розраховується як відношення максимального та мінімального значення Іп КМАФАНМ до RMSEP моделі.

$$RMSEP = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{N}}, \quad (5)$$

де y_i та $\hat{y}_i$ – вихідні та прогнозовані значення Іп КМАФАНМ, відповідно; N – кількість тестових зразків.

Таблиця 4. Значення зважених регресійних коефіцієнтів X -змінних PLS моделей прогнозування мікробіологічного псування ковбас в.с.

Параметри електронного носа	b	Параметри електронного носа	b
PLS1, Factor-1, b_0 : 13.2055			
$\Delta F^{\max}$ PEGseb	-0.0841	S PEGseb	-0.0850
$\Delta F^{\max}$ PEGad	-0.0725	S PEGad	-0.0663
$\Delta F^{\max}$ DCG18k6	0.0907	S DCG18k6	0.1173
$\Delta F^{\max}$ TX-100	0.1057	S TX-100	0.0947
$\Delta F^{\max}$ PDEGsuc	0.0851	S PEG-2000	-0.0808
$\Delta F^{\max}$ PVP	-0.1113	S PDEGsuc	-0.1096
		S PVP	-0.1160
PLS2, Factor-1, b_0 : 20.3403			
S PEGseb	-0.1509	S PEG-2000	-0.1434
S PEGad	-0.1177	S PDEGsuc	-0.1947
S DCG18k6	0.2082	S PVP	-0.2060
S TX-100	0.1682		
PLS3, Factor-3, b_0 : -1.7361			
$\Delta F^{\max}$ PEGseb	-0.1857	$\Delta F^{\max}$ TX-100	0.3002
$\Delta F^{\max}$ PEGad	0.2333	$\Delta F^{\max}$ PDEGsuc	0.3402
$\Delta F^{\max}$ DCG18k6	0.1611	$\Delta F^{\max}$ PVP	-0.3788

Значення RMSE cal, RMSE val та RMSEP в програмі виражається в вихідних одиницях Y, тобто в одиницях ln КМАФАнМ.

Арбітражний метод визначення загального числа мікробів в 1 г продукту не має метрологічних характеристик [17]. За кінцевий результат підрахунку кількості КМАФАнМ (КУО/г) приймають середнє арифметичне значення результатів двох розведень різної маси, таким чином точність аналізу залежить лише від кваліфікації та уважності аналітика.

Для надання метрологічних характеристик нового методу розраховували також значення середньої відносної похибки прогнозування ln КМАФАнМ зразків тестової вибірки. Порівнювали точність апроксимації трьох калібрувальних моделей та придатність їх для прогнозування (табл. 5).

Перевірка надійності моделей для прогнозування загального мікробного числа базується на визначенні схожості нового зразка до калібрувальних зразків. На графіках (рис. 2) прогнозоване значення – це червона горизонтальна лінія з червоною вертикальною лінією RMSEP, а синій прямокутник – відхилення кожного зразка. Відхилення вказує наскільки надійними є прогнозовані результати, оскільки враховує залишок та коефіцієнт впливу спостереження відносно калібрувальних даних.

Встановлено, що PLS1 та PLS2 регресійні моделі достатньо точно прогнозують загальне мікробне число ln КМАФАнМ (табл. 5, рис. 2). Велика похибка моделювання PLS3 та прогнозування нових зразків вказує на те, що параметри $\Delta F_i^{\max}$ містять недостатньо інформації про динаміку зміни газової фази в процесі псування ковбас. Додавання $\Delta F_i^{\max}$ параметрів до змінних S_i призводить до незначного погіршення точності прогнозування моделі PLS1, однак під час моделювання рекомендують використовувати найменшу оптимальну кількість змінних для уникнення нестабільності предикторів [29].

Як видно з графіків прогнозування (рис. 2), найменша помилка RMSEP для регресійної моделі PLS2 за зміни ln КМАФАнМ в межах від 5.44 до 9.20. Значення нових зразків збігаються з калібрувальним набором в межах розрахованого відхилення. Більша дисперсія калібрувальних даних для A33, A34 та A43 тестових зразків пояснюється мінливістю газової фази зразків на межі норми показника КМАФАнМ=1.0...1.1·10³ КУО/г, що не становить проблеми для практичного використання методики.

Для розрахування метрологічних характеристик методу в одиницях КУО/г показника КМАФАнМ ковбасних виробів застосовували потенціювання прогнозованих значень тестового набору зразків (табл. 6).

Таблиця 5. Точність моделювання регресійних моделей та прогнозування ln КМАФАнМ.

Метрологічні характеристики методики	PLS1	PLS2	PLS3
Точність регресійних моделей			
RMSE cal	0.31	0.31	0.35
RMSE val	0.34	0.34	0.44
R ² cal	0.95	0.95	0.94
R ² val	0.94	0.94	0.90
Точність прогнозування нових зразків			
RMSEP	0.19	0.12	0.39
Відносна похибка RMSEP, %	2.0...3.5	1.3...2.3	4.2...7.2
Відносна похибка вимірювання, %	2.1	1.4	4.46

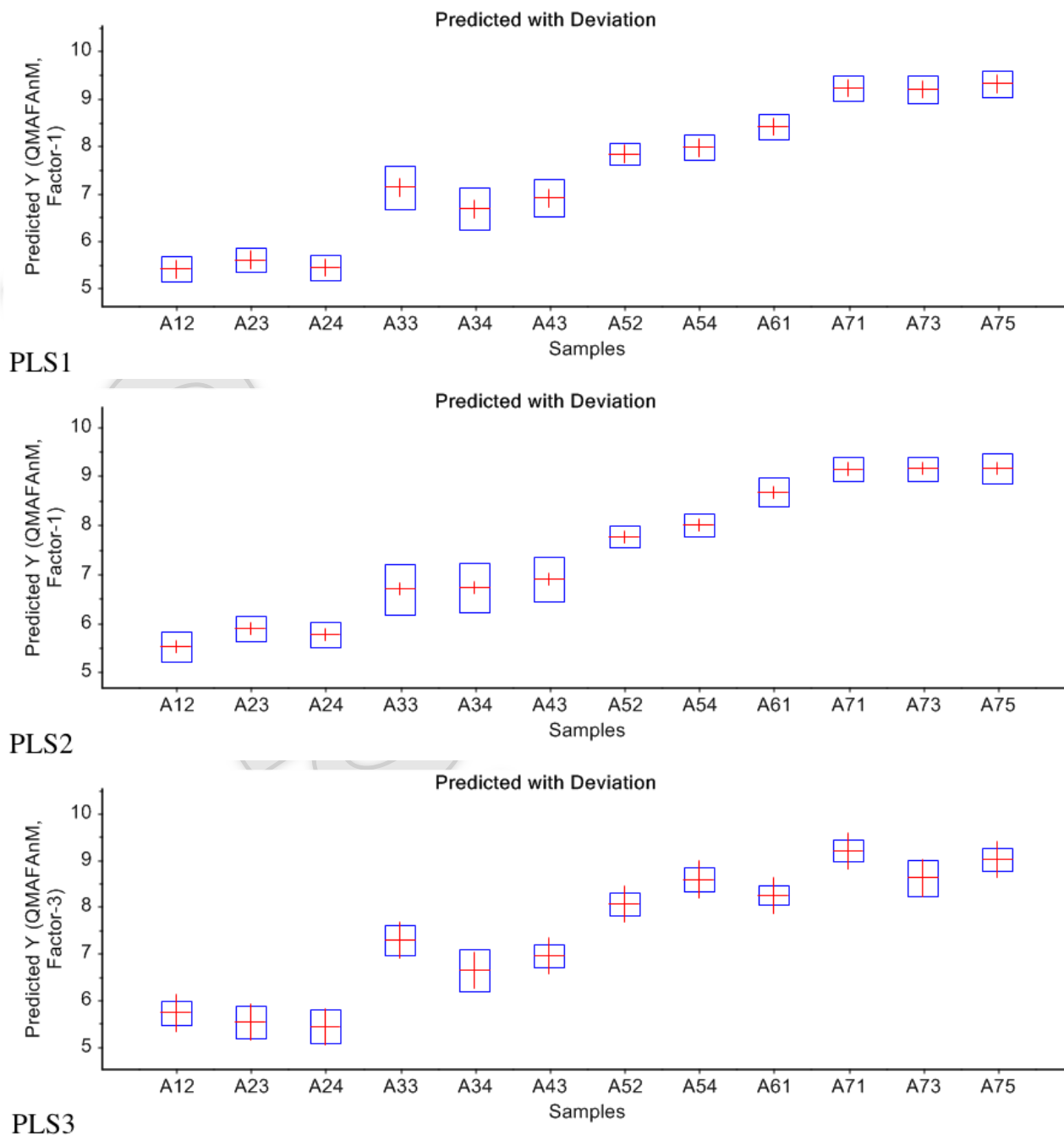


Рис.2. Результати прогнозування Іп КМАФАНМ тестових зразків ковбасних виробів (в.с.).

Таблиця 6. Метрологічні характеристики методу прогнозування показника КМАФАНМ ковбас в.с. із використанням параметрів S_e електронного носа та PLS регресії.

Значення КМАФАНМ, КУО/г	Відносна похибка вимірювання, %
від $2.3 \cdot 10^2$ до $9.9 \cdot 10^3$	10.0
не більше ніж $1.0 \cdot 10^3$	11.6
$1.1 \cdot 10^3$ та більше	8.4

Регресійна модель, побудована з використанням параметрів S_i електронного носа та методу проєкції на латентні структури дозволяє зробити надійний прогноз мікробіологічного показника КМАФАнМ та оцінити безпечність ковбасного виробу в.с. з відносною похибкою вимірювання 11.6% для зразків з кількістю мікроорганізмів в межах норми стандарту [11].

На основі наведеного вище дослідження, використовували S_i параметри сенсорів як X-змінні для побудови моделі прогнозування $\ln$ КМАФАнМ ковбасним виробів із вмістом соєвого ізоляту (10-20 мас.%). Одновимірний кореляційний аналіз застосовували для формування матриці предикторів, що суттєво корелюють із прогнозованим мікробіологічним показником. З X-матриці предикторів вилучили параметр S_{PVP} , значення якого нижче критичного значення коефіцієнта кореляції Пірсона для набору даних $r_{72, 0.05} = 0.23$ (табл. 7).

Таблиця 7. Значення коефіцієнтів кореляції параметрів електронного носа з мікробіологічним показником КМАФАнМ ковбас з СІ.

Параметри відгуків сенсорів	r	Параметри відгуків сенсорів	r
S PEGseb	-0.68	S PEG-2000	0.85
S PEGad	-0.55	S PDEGsuc	0.75
S DCG18k6	0.61	S PVP	-0.06
S TX-100	0.63		

Необхідність будівництва окремих прогностичних моделей для м'ясо-соевих та м'ясних зразків підтверджують результати кореляційного аналізу (табл.3, табл.7). Для м'ясо-соевих виробів характерна висока ступінь кореляції показника $\ln$ КМАФАнМ з параметрами сенсорів з покриттями поліетиленгліколь 2000 та полідиетиленгліколь сукцинат, в той час як для м'ясних ковбас в.с. найбільш значущі параметри сенсорів з покриттями полівінілпіролідон та дициклогексано-18-краун-6, а параметр сенсора поліетиленгліколь 2000 навіть вилучили з матриці предикторів у зв'язку з відсутністю лінійної залежності між змінними.

Навчальна вибірка для побудови калібрувальної моделі складалася з 54 зразків, тобто 27 зразків кожного типу ковбас із вмістом 10% мас. та 20% мас. СІ. Калібрувальну модель перевіряли за допомогою повної перехресної перевірки з числом сегментів 54 (leave-one-out). Матриці даних X та Y автошкалували ($1/Sdev$) для отримання одиничної дисперсії всіх змінних.

Оцінювали прогностичні можливості побудованої регресійної моделі за тестовим набором з 18 нових зразків.

Для PLS моделі прогнозування ковбасних виробів із СІ використовували 2 ГК з 75% X та 91% Y пояснювальними дисперсіями. Обрано мінімальну кількість факторів, що описують максимальний відсоток Y дисперсії.

Представлені дані на графіку рахунків наочно показують (рис. 3) наскільки різні за складом газової фази є зразки В (10 мас.% СІ) та С (20 мас.% СІ).

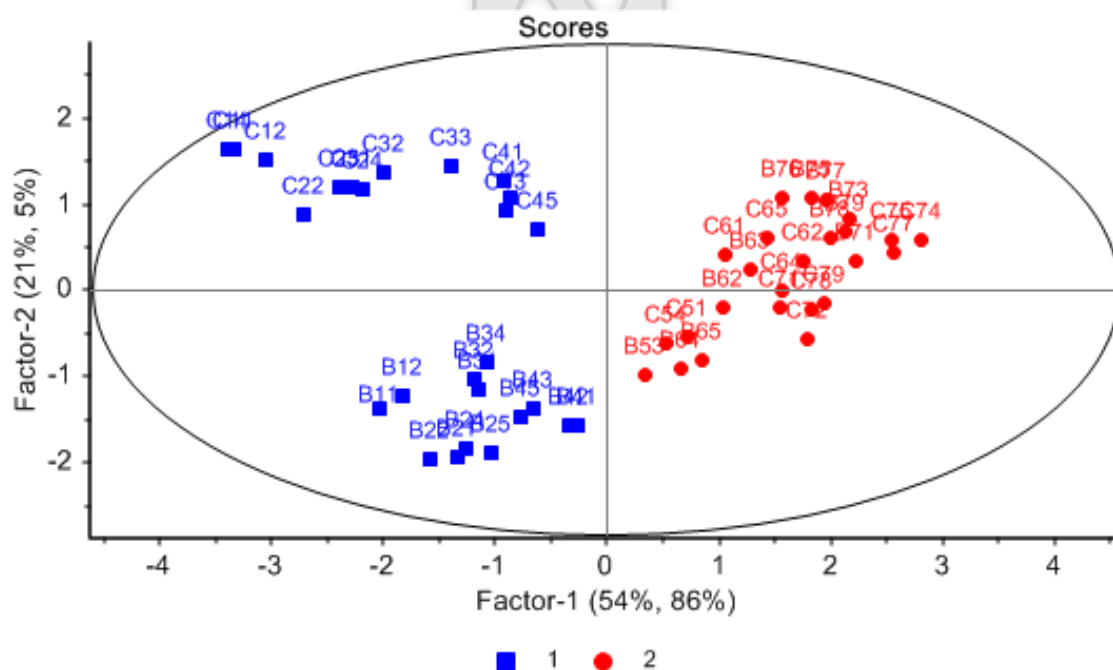


Рис. 3. Графік рахунків PLS моделі прогнозування $\ln$ КМАФАнМ ковбас з 10% мас. СІ (В) та 20% мас. СІ (С).

Зразки з загальним мікробним числом в межах норми стандарту (рис.3, позначення 1) розподіляються на дві групи. В процесі мікробіологічного псування ковбаси В та С стають більш схожі за складом летких сполук (рис.3, позначення 2, показник КМАФАнМ за межами норми), однак, розташування точок на графіку рахунків вказує на різний характер перебігу мікробіологічного псування ковбас (нелінійність зразків В). ГК1 описує процес мікробіологічного псування, в той час як ГК2 описує різницю між зразками із 10% мас. та 20% мас. СІ. Спираючись на дані, що складаються з двох груп зразків, може бути складно побудувати загальну прогностичну модель. В такому випадку модель стає узагальненою, вона може визначити, наприклад, до якої групи відноситься новий зразок, але точність всередині групи може бути низькою.

Рівняння прогнозування показника $\ln$ КМАФАнМ для м'ясо-соевих ковбасних виробів, з урахуванням двох ГК має вигляд:

$$\hat{y} = 1 \times (-1.9832) + X\hat{b}, \quad (6)$$

де 1 – одинична матриця, X – матриця предикторів (параметрів електронного носа), $\hat{b}$ – вектор зважених регресійних коефіцієнтів, значення яких наведено в табл. 8.

Встановлено, що найбільш вагомими для моделі є параметри сенсорів з покриттями поліетиленгліколь адипінат та поліетиленгліколь себацінат (від'ємна кореляція з показником $\ln$ КМАФАнМ), поліетиленгліколь 2000 та полідиетиленгліколь сукцинат (позитивна кореляція з показником $\ln$ КМАФАнМ), що проявляють високу масову чутливість до парів спиртів,

альдегідів, кислот та амінів [1,28], метаболітів мікроорганізмів псування, а також специфічних мікроорганізмів соєвих продуктів. М'ясо-соеві вироби характеризуються більшою швидкістю зростання кількості мікроорганізмів псування та їх метаболічною активністю.

Таблиця 8. Значення зважених регресійних коефіцієнтів PLS моделі прогнозування мікробіологічного псування м'ясо-соевих ковбас.

Параметри електронного носа	b	Параметри електронного носа	b
S PEGseb	-0.2140	S TX-100	0.0945
S PEGad	-0.2883	S PEG-2000	0.2645
S DCG18k6	0.1114	S PDEGsuc	0.3563

На графіку кореляції навантажень X та Y відмічені предиктори з найбільш вагомими регресійними коефіцієнтами (рис.4). Аналіз графіку вказує на те, що всі предиктори корелюють з прогнозованим показником та є значущими для моделі, так як знаходяться між двома колами та далеко від центру координат.

Результати оцінки моделювання регресійної моделі та її прогностичних можливостей наведено в табл.9. Для двох ГК величина RMSEP складала 0.41 за зміни $\ln$ КМАФАнМ в межах від 5.60 до 10.24, що дає середню відносну помилку біля 5%. На графіку порівняння прогнозованих та вихідних значень Y кореляція становить 0.97, та дані лежать близько до регресійної прямої, що підтверджує хорошу точність моделі. Так як помилка в референтному методі невідома, складно оцінити точність нового методу в порівнянні зі стандартним.

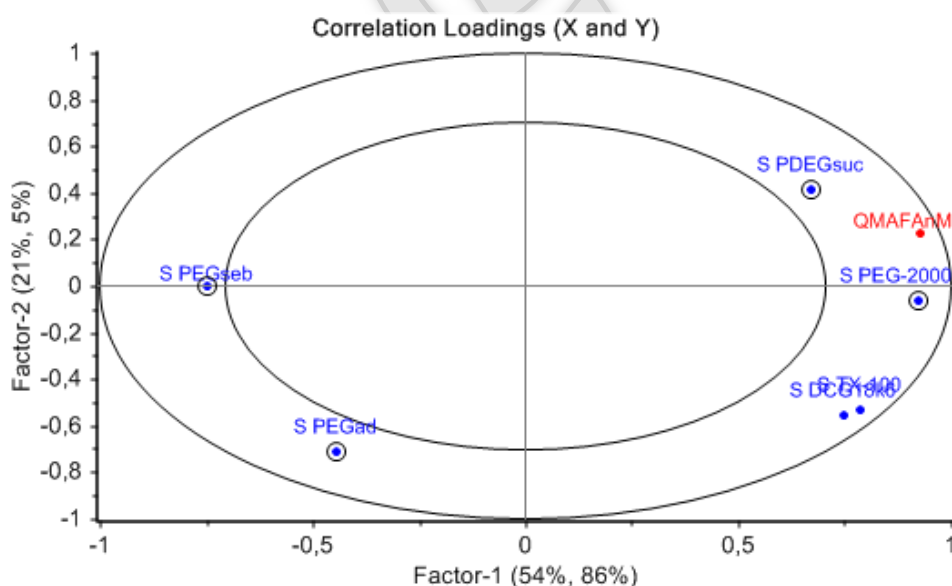


Рис. 4. Графік кореляції навантажень X та Y змінних PLS моделі прогнозування $\ln$ КМАФАнМ ковбас із СІ.

Таблиця 9. Точність моделювання регресійної моделі та прогнозування ln КМАФАнМ м'ясо-соевих ковбас.

Метрологічні характеристики методики	PLS
Точність моделювання	
RMSE cal	0.49
RMSE val	0.52
R ² cal	0.92
R ² val	0.91
Оцінка моделі за новими зразками	
RMSEP	0.41
Відносна похибка RMSEP, %	4.0...7.4
Відносна похибка вимірювання, %	2.8

Результати перевірки надійності PLS моделі для прогнозування загального мікробного числа 18 нових зразків ковбасних виробів з 10 мас. % CI (B) та 20 мас. % CI (C) наведені на рис. 5. Зразки B72 та C 73 можуть бути викидами, так як прогнозовані значення занадто великі для моделі. Показник КМАФАнМ зазначених зразків дорівнює $2.2 \cdot 10^4$ та $2.8 \cdot 10^4$ КУО/г відповідно, що є далеко за межами норм стандарту, прогнозування таких зразків під час практичного використання методики є малоцікаво. Зразок B54 має занадто низьке прогнозоване значення для моделі, однак, не відрізняється від зразків калібрувального набору, що може вказувати на недостатню калібрувальну вибірку зразків з показником КМАФАнМ на межі норми.

Для розрахування метрологічних характеристик методу в одиницях КУО/г показника КМАФАнМ м'ясо-соевих ковбасних виробів застосовували потенціювання прогнозованих значень тестового набору зразків.

Таблиця 10. Метрологічні характеристики методу прогнозування показника КМАФАнМ м'ясо-соевих ковбас із використанням параметрів S_i електронного носа та PLS регресії.

Значення КМАФАнМ, КУО/г	Відносна похибка вимірювання, %
від $2.7 \cdot 10^2$ до $2.8 \cdot 10^4$	19.7
не більше ніж $2.5 \cdot 10^3$	11.7
$2.6 \cdot 10^3$ та більше	29.8

Отже, побудована регресійна модель, незважаючи на неоднорідність вибірки та відмінності газової фази двох груп м'ясо-соевих ковбас (10-20 % мас. CI), є достатньо надійною для прогнозування мікробіологічного показника зразків з КМАФАнМ не більше ніж $2.5 \cdot 10^3$ КУО/г, відносна похибка вимірювання становить 11.7 %.

Розроблені моделі прогнозування показника КМАФАнМ ковбасних виробів із використанням параметрів S_i електронного носа в поєднанні з методом проекції на латентні структури дозволять провести експрес-аналіз мікробіологічного псування ковбас з відносною похибкою вимірювання менше 12 %, що є достатньою точністю для скринінгу об'єктів. Впровадження даної методики дозволить суттєво зменшити тривалість аналізу та знизити собівартість, уникнути суб'єктивності оцінки результатів.

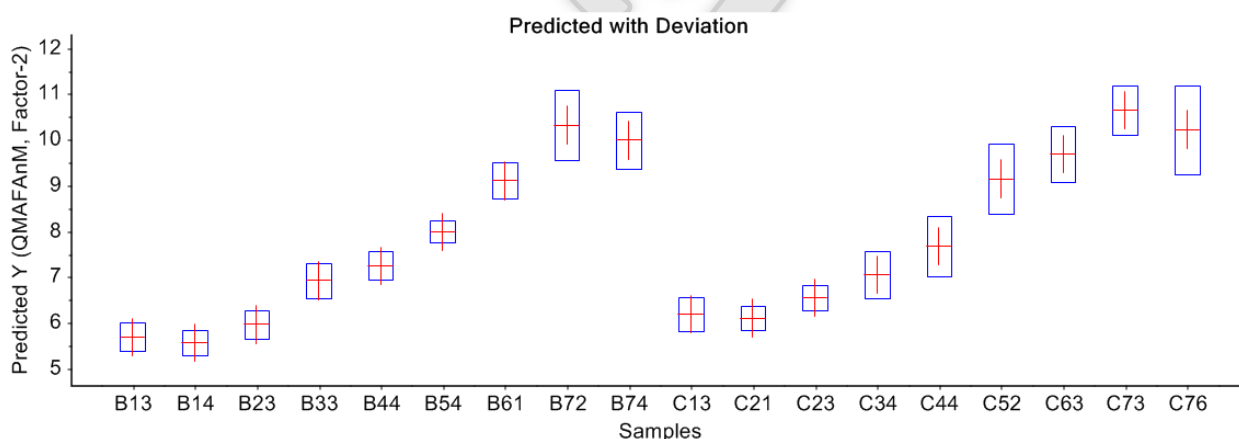


Рис. 5. Результати прогнозування ln КМАФАнМ тестових зразків ковбасних виробів з 10 % мас. CI (B) та 20 % мас. CI (C).

Розроблення схеми інтелектуального аналізу хімічних образів електронного носу для розв'язання задач якісного та кількісного аналізу

Отримані результати свідчать про те, що інтелектуальна система розпізнавання образів запахів дозволяє оцінити різні характеристики об'єкта за одне вимірювання.

Надмірність масиву даних електронного носу усувається за рахунок вилучення інформативних параметрів з вихідних відгуків сенсорів відповідно до задачі аналізу, параметричної оптимізації алгоритмів обробки багатовимірних даних під час їх синтезу. За відсутності інформації про взаємозв'язок параметрів з необхідними властивостями об'єкта пошук оптимальних параметрів проводили експериментальним шляхом, ітераційно оптимізуючи алгоритми для

кращого описання даних.

Для вирішення задач якісного аналізу (багатокласової класифікації ковбас, ідентифікації, дискримінації ковбас різного складу) та кількісного аналізу (прогнозування показника загального мікробного числа) запропоновано схему інтелектуального аналізу даних електронного носу, що узагальнює проведені вище дослідження (рис. 6).

Використання наведеної схеми дозволяє з одного багатовимірного відклику електронного носу оцінити автентичність зразків, виявити фальсифікацію та прогнозувати безпечність харчового продукту. Аналітичний контроль ковбасних виробів може бути ефективно проведений з використанням запропонованих класифікаційних та регресійних моделей.

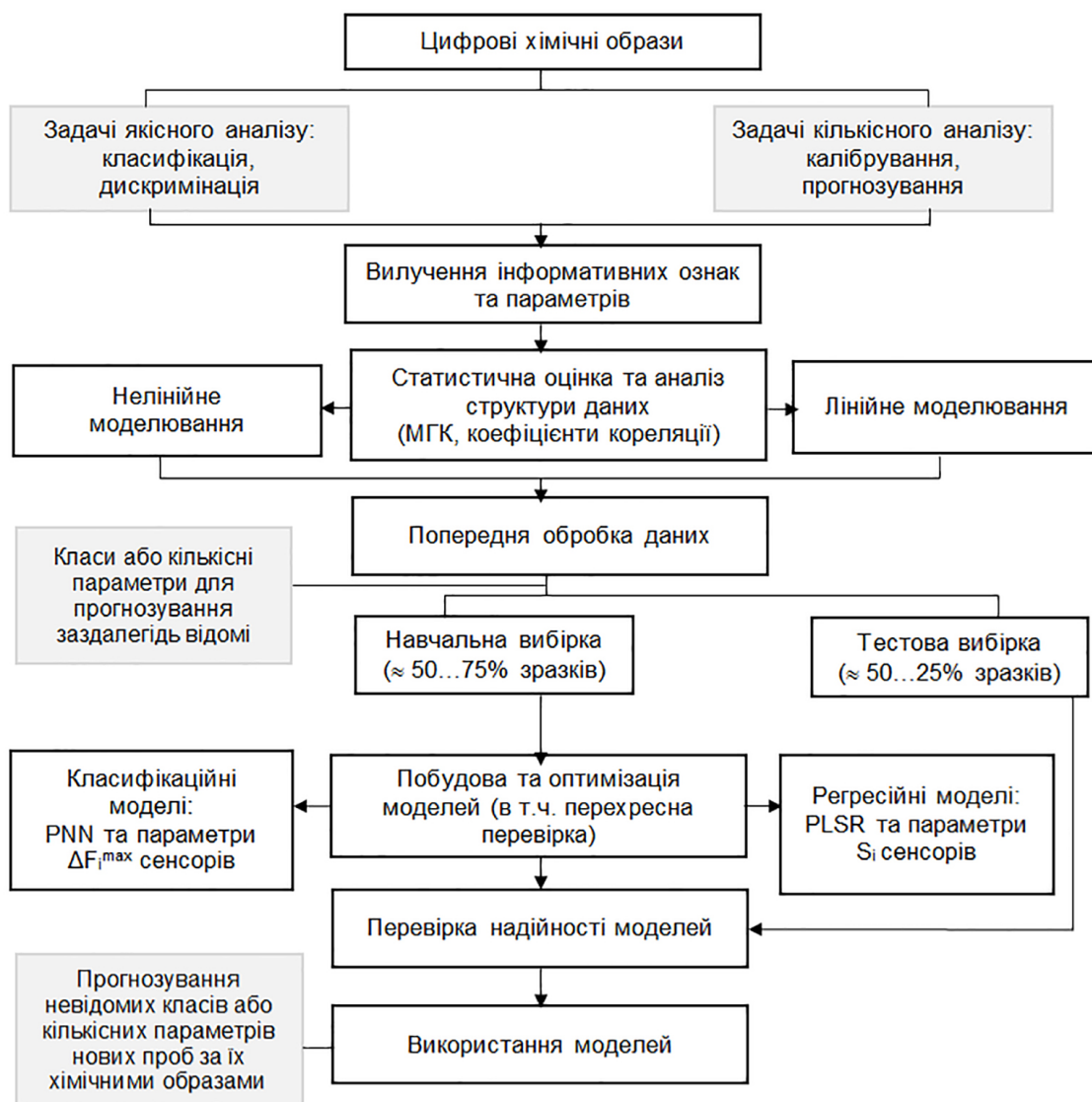


Рис. 6. Схема інтелектуального аналізу хімічних образів мультисенсорної системи для класифікації зразків (ідентифікації, виявлення фальсифікації) та прогнозування кількісних параметрів (кількості мікроорганізмів, вмісту заміника).

Висновки

Запропоновано методику інтелектуального аналізу багатовимірних даних електронного носа для оцінки автентичності ковбасних виробів, виявлення фальсифікації соєвим ізолятом та визначення показника загального мікробного числа за одне вимірювання.

Показано, що для розпізнавання хімічних образів запахів для розв'язання задачі ідентифікації ковбасних виробів кращу точність має класифікаційна модель з використанням параметрів максимальних сигналів сенсорів $\Delta F_i^{\max}$, отже, стадія адсорбції містить інформативні дискримінаційні ознаки.

Інтегральні параметри S_i краще відображають зміни загальної кількості метаболітів протягом зберігання ковбасних виробів різного складу, в даному випадку фаза десорбції є більш інформативною, ніж адсорбції, даний параметр рекомендується для будування регресійних моделей прогнозування показника безпечності.

Розроблено спосіб оцінки автентичності

ковбасних виробів різних сортів та виявлення фальсифікації соєвим ізолятом із використанням максимальних сигналів сенсорів електронного носа та імовірнісної нейронної мережі, 100 % точність прогнозування зразків отримано за значень параметра згладжування мережі $\sigma = 0.2 - 1.0$.

Запропоновано спосіб прогнозування загального мікробного числа (КМАФАнМ, КУО/г) ковбасних виробів із використанням оптимізованих матриць параметрів площ під кривими відгуків сенсорів та методу проєкції на латентні структури. Отримані надійні регресійні моделі з використанням трьох значущих факторів у випадку ковбас в. с. та двох ГК для ковбас, що містять СІ в складі, що мають відносну помилку вимірювання менше 12 %

Запропонована методика з використанням мультисенсорної системи в поєднанні з алгоритмами машинним навчанням дозволить провести аналітичний контроль ковбасних виробів та суттєво зменшити тривалість аналізу в порівнянні із стандартними методами, знизити собівартість аналізу, уникнути суб'єктивності оцінки результатів.

Література

1. Kuchmenko, T. A.; Pogrebnya, D. A.; Shuba, A. A. Application of Principal Component Analysis and Discriminant Analysis Using Latent Structure Regression to the Identification of Highly Volatile Organic Compounds in Gas Mixtures Based on Piezoelectric Quartz Microweighing Parameters. *J. Anal. Chem.* **2013**, 68(7), 652–661. doi:10.1134/s106193481307006x.
2. Di Rosa, A. R.; Leone, F.; Cheli, F.; et al. Fusion of Electronic Nose, Electronic Tongue and Computer Vision for Animal Source Food Authentication and Quality Assessment – A Review. *J. Food Eng.* **2017**, 210, 62–75. doi: 10.1016/j.jfoodeng.2017.04.024.
3. Ropodi, A. I.; Panagou, E. Z.; Nychas, G. J. E. Data Mining Derived from Food Analyses Using Non-Invasive/Non-Destructive Analytical Techniques; Determination of Food Authenticity, Quality & Safety in Tandem with Computer Science Disciplines. *Trends in Food Science and Technology*. Elsevier Ltd April 1, 2016, pp 11–25. doi: 10.1016/j.tifs.2016.01.011.
4. Ghasemi-Varnamkhasti, M.; Apetrei, C.; Lozano, J.; et al. Potential Use of Electronic Noses, Electronic Tongues and Biosensors as Multisensor Systems for Spoilage Examination in Foods. *Trends Food Sci. Technol.* **2018**, 80, 71–92. doi: 10.1016/j.tifs.2018.07.018.
5. Yan, J.; Guo, X.; Duan, S.; et al. Electronic Nose Feature Extraction Methods: A Review. *Sensors* **2015**, 15 (11), 27804–27831. doi:10.3390/s151127804.
6. Kalinichenko, A.; Arsenyeva, L.; Pasichnij, V. Використання Електронного Носа Та Імовірнісної Нейронної Мережі Для Ідентифікації Ковбас.

Visnik kiivskogo nacionalnogo universitetu imeni T. Shevchenka **2017**, 2 (54), 47–51. (in Ukr.)

7. Arsenyeva, L. Y.; Kalinichenko, A. A.; Ivanov, S. V.; et al. Method for Determining the Soy Content in Sausages. Pat. 111254 Ukraine, MPK G01N 33/02 MPK A23L 13/00., 2015. (in Ukr.)
8. Arsenyeva, L. Y.; Kalinichenko, A. A.; Ivanov, S. V.; et al. Method of Detecting the Adulteration of Sausages by Soy Products. Pat. 111253 Ukraine, MPK G01N 33/02 MPK A23L 13/60, 2015. (in Ukr.)
9. Abbas, O.; Zdravec, M.; Baeten, V.; et al. Analytical Methods Used for the Authentication of Food of Animal Origin. *Food Chem.* **2018**, 246, 6–17. doi: 10.1016/j.foodchem.2017.11.007.
10. Peris, M.; Escuder-Gilabert, L. Electronic Noses and Tongues to Assess Food Authenticity and Adulteration. *Trends Food Sci. Technol.* **2016**, 58, 40–54. doi: 10.1016/j.tifs.2016.10.014.
11. DSTU 4436:2005. *Boiled Sausages, Frankfurters, Meat Loaves*; Ukraine, 2006; p 32. (in Ukr.)
12. Horwitz, W.; Latimer, G. W. *Official Methods of Analysis of AOAC International*; 2005; pp 1–24.
13. Montowska, M.; Fornal, E. Detection of Peptide Markers of Soy, Milk and Egg White Allergenic Proteins in Poultry Products by LC-Q-TOF-MS/MS. *LWT - Food Sci. Technol.* **2018**, 87, 310–317. doi:10.1016/j.lwt.2017.08.091
14. Hoffmann, B.; Münch, S.; Schwägele, F.; et al. A Sensitive HPLC-MS/MS Screening Method for the Simultaneous Detection of Lupine, Pea, and Soy Proteins in Meat Products. *Food Control* **2017**, 71, 200–209. doi:10.1016/j.foodcont.2016.06.021.
15. Ducatti, R.; de Almeida Nogueira Pinto, J. P.; Sartori, M. M. P.; et al. Quantification of Soy Protein

Using the Isotope Method ($\Delta^{13}\text{C}$ and $\Delta^{15}\text{N}$) for Commercial Brands of Beef Hamburger. *Meat Sci.* **2016**, 122, 97–100. doi:10.1016/j.meatsci.2016.07.012.

16. DSTU 8379:2015. *Meat and Meat Products. Accelerated Method for Determining the Composition of Raw Products*; TK 140: Ukraine, 2017; p 13. (in Ukr.)

17. DSTU 8720:2017. *Sausage Products and Meat Products. Methods for Determining Microbial Contamination*; TK 140: Ukraine, 2019; p 64. (in Ukr.)

18. Qiu, S.; Wang, J. The Prediction of Food Additives in the Fruit Juice Based on Electronic Nose with Chemometrics. *Food Chem.* **2017**, 230, 208–214. doi:10.1016/j.foodchem.2017.03.011.

19. Papadopoulou, O. S.; Panagou, E. Z.; Mohareb, F. R.; et al. Sensory and Microbiological Quality Assessment of Beef Fillets Using a Portable Electronic Nose in Tandem with Support Vector Machine Analysis. *Food Res. Int.* **2013**, 50 (1), 241–249.

20. Falasconi, M.; Comini, E.; Concina, I.; et al. Electronic Nose and Its Application to Microbiological Food Spoilage Screening. In *Sensing Technology: Current Status and Future Trends II*; Springer International Publishing Switzerland, 2014; pp 119–140.

21. Esbensen, K. H.; Guyot, D.; Houmøller, L. P. *Multivariate Data Analysis – In Practice*, 5 Edition.; Camo Process AS: Oslo, 2004.

22. Demuth, H.; Beale, M. *Neural Network Toolbox User's Guide*; The MathWorks Inc., 2002.

23. Ríos, A.; Barceló, D.; Buydens, L.; et al. Quality Assurance of Qualitative Analysis in the Framework of the European Project “MEQUALAN.” *Accredit. Qual. Assur.* **2003**, 8 (2), 68–77.

24. Specht, D. Probabilistic Neural Networks. *Neural Networks* **1990**, 3 (1), 109–118. doi:10.1016/0893-6080(90)90049-Q.

25. *Food Spoilage Microorganisms*; Blackburn, C. de W., Ed.; Woodhead Publishing, 2006.

26. Solina, M.; Baumgartner, P.; Johnson, R. L.; et al. Volatile Aroma Components of Soy Protein Isolate and Acid-Hydrolysed Vegetable Protein. *Food Chem.* **2005**, 90 (4), 861–873. doi:10.1016/j.foodchem.2004.06.005.

27. Fincheira, P.; Quiroz, A. Microbial Volatiles as Plant Growth Inducers. *Microbiol. Res.* **2018**, 208 (January), 63–75. doi:10.1016/j.micres.2018.01.002.

28. Kuchmenko, T. A.; Umarmkhanov, R. U. Peculiarities of Microweighing of Trace Quantities of Alkylamines on Polymer and Solid-State Thin Films. *J. Anal. Chem.* **2013**, 68 (4), 368–375. doi:10.1134/s1061934813040059.

29. Hoskuldsson, A. PLS REGRESSION METHODS. *J. Chemom.* **1988**, 2 (August 1987), 211–228.