

ISSN 2220-637X

Вісник
Харківського
Національного
Університету
імені В. Н. Каразіна

СЕРІЯ «ХІМІЯ»
Вип. 38 (61)

Kharkiv University Bulletin
Chemical series. Issue 38 (61)

Заснований 1935 року як
“Труди інституту хемії при
Харківському державному
університеті”

Published since 1935; initially
under the title "Proceedings of
the Institute of Chemistry at
Kharkiv State University"

Харків Kharkiv
2022

Вісник містить статті, присвячені різним аспектам теоретичної хімії, хімічного аналізу, органічної хімії, спектроскопії, фізико-хімії розчинів та поверхневих явищ, електрохімії, хімічного матеріалознавства.

Для науковців і фахівців. Видання є фаховим в галузі хімічних наук.
(Наказ Міністерства освіти і науки України № 1643 від 28.12.2019 року)

Затверджено до друку рішенням Вченої ради Харківського національного університету імені В.Н. Каразіна (протокол № 15 від 26 вересня 2022 р.)

Головний редактор

О.І. Коробов

д.х.н., проф., ХНУ імені В.Н. Каразіна, Україна

Редактори

А.О. Дорошенко

д.х.н., проф., ХНУ імені В.Н. Каразіна, Україна

М.О. Мчедлов-Петросян

д.х.н., проф., ХНУ імені В.Н. Каразіна, Україна,
член-кореспондент НАН України

Технічний редактор

А.Б. Захаров

к.х.н., ХНУ імені В.Н. Каразіна, Україна

Редакційна рада

В.В. Іванов

д.х.н., проф., ХНУ імені В.Н. Каразіна, Україна

В.О. Черановський

д.ф.-м.н., проф., ХНУ імені В.Н. Каразіна, Україна

С.А. Шаповалов

д.х.н., проф., ХНУ імені В.Н. Каразіна, Україна

О.І. Юрченко

д.х.н., проф., ХНУ імені В.Н. Каразіна, Україна

Міжнародна консультативна рада

О.М. Калугін

к.х.н., проф., ХНУ імені В.Н. Каразіна, Україна

А.Ю. Назаренко

PhD, Prof., Buffalo State College, USA

О.В. Преждо

PhD, Prof., University of Southern California, USA

В.А. Чебанов

д.х.н., проф., НТК «Інститут монокристалів», Україна,
член-кореспондент НАН України

Редактори консультанти

І.М. В'юник

д.х.н., проф., ХНУ імені В.Н. Каразіна, Україна

В.І. Ларін

д.х.н., проф., ХНУ імені В.Н. Каразіна, Україна

В.І. Лебідь

д.х.н., проф., ХНУ імені В.Н. Каразіна, Україна

Адреса редакційної колегії: Україна, 61022, Харків, майдан Свободи, 4,
ХНУ імені В.Н. Каразіна, хімічний факультет; тел.: +38 057 707 51 29.

E-mail: chembull@karazin.ua a.korobov@karazin.ua

Статті пройшли внутрішнє та зовнішнє рецензування.

Свідоцтво про державну реєстрацію КВ № 21563-11463Р від 27.07.2015.

© Харківський національний університет
імені В.Н. Каразіна, оформлення, 2022

The Chemical Series publishes papers devoted to various aspects of theoretical chemistry, chemical analysis, organic chemistry, inorganic chemistry, physical chemistry of solutions and surface phenomena, electrochemistry, materials chemistry. The bulletin is officially authorized by the Highest Attestation Commission of Ukraine to publish results of research submitted for PhD and ScD degrees.

Publication of this issue is approved by the Academic Council of V.N. Karazin Kharkiv National University (protocol № 15 from 26.09.2022).

Editor-in-chief

Alexander Korobov ScD, Prof., V.N. Karazin Kharkiv National University

Editors

Andrey Doroshenko ScD, Prof., V.N. Karazin Kharkiv National University
Nikolay Mchedlov-Petrosyan ScD, Prof., V.N. Karazin Kharkiv National University, corresponding member of NAS of Ukraine

Managing editor

Anton Zakharov PhD, V.N. Karazin Kharkiv National University

Editorial board

Vladimir Ivanov ScD, Prof., V.N. Karazin Kharkiv National University
Vladislav Cheranovsky ScD, Prof., V.N. Karazin Kharkiv National University
Sergey Shapovalov ScD, Prof., V.N. Karazin Kharkiv National University
Oleg Yurchenko ScD, Prof., V.N. Karazin Kharkiv National University

International advisory board

Oleg Kalugin PhD, Prof., V.N. Karazin Kharkiv National University
Alexander Nazarenko PhD, Prof., Buffalo State College, USA
Oleg Prezhdo PhD, Prof., University of Southern California, USA
Valentyn Chebanov ScD, Prof., SSI "Institute for Single Crystals", Ukraine corresponding member of NAS of Ukraine

Consulting editors

Ivan Vyunnik ScD, Prof., V.N. Karazin Kharkiv National University
Vasyl Larin ScD, Prof., V.N. Karazin Kharkiv National University
Valentyn Lebed ScD, Prof., V.N. Karazin Kharkiv National University

Address of editorial team: 4 Svobody Sq., Kharkiv, 61022, Ukraine, V.N. Karazin Kharkiv National University, School of Chemistry; tel.: +38 057 707 51 29.
E-mail: chembull@karazin.ua a.korobov@karazin.ua

All articles have been reviewed.

Certificate of state registration KB № 21563-11463P from 27.07.2015.

© V.N. Karazin Kharkiv National University,
Design, 2022

ЗМІСТ

- 6 Робастне оцінювання параметрів регресій. Теорія нечіткості та інші моделі.
А.В. Пантелеймонов, Д.О. Анохін, В.В. Іванов
- 16 Дослідження змішаних міцел додецилсульфату натрію – 1-бутанолу методом молекулярної динаміки. ***І.В. Кононенко, В.С. Фарафонов***
- 23 Особливості сольватохромії 4-[[[(2,4-динітрофеніл)метилен]іміно-2,6-дифеніл]фенола та барвника Райхардта. DFT розрахунки. ***Є. О. Сергєєва, А. Б. Захаров, С. М. Кійко***

CONTENTS

- 6 Robust evaluation of regression parameters. The fuzzy theory and other models.
A.V. Panteleimonov, D.O. Anokhin, V.V. Ivanov
- 15 Examination of mixed sodium dodecyl sulfate – 1-butanol micelles with molecular dynamics.
I.V. Kononenko, V.S. Farafonov
- 23 Peculiarities of solvatochromism of 4-[[[(2,4-dinitrophenyl)methylene]imino-2,6-diphenyl]phenol
and Reichardt's dye. DFT calculations. ***Y.O. Serhieieva, A.B. Zakharov, S.M. Kiyko***

РОБАСТНЕ ОЦІНЮВАННЯ ПАРАМЕТРІВ РЕГРЕСІЙ. ТЕОРІЯ НЕЧІТКОСТІ ТА ІНШІ МОДЕЛІ

А. В. Пантелеймонов^а, Д. О. Анохін^б, В. В. Іванов^с


Харківський національний університет імені В. Н. Каразіна, майдан Свободи 4, Харків, 61022, Україна

a) ✉ panteleimonov@karazin.ua

b) ✉ dmitriy25102002@gmail.com

c) ✉ vivanov@karazin.ua

 <https://orcid.org/0000-0003-0265-1264>

 <https://orcid.org/0000-0002-4958-2692>

 <https://orcid.org/0000-0003-2297-9048>

Представлене співставлення результатів розрахунків параметрів лінійної регресії на основі теорії нечіткості та інших статистичних підходів. Запропоновано алгоритм простого зваженого методу найменших квадратів, що не спирається на апріорну інформацію щодо розподілу похибок вимірювань. Роботу алгоритму перевірено на модельних даних, його адекватність підтверджено на основі застосування широковживаних критеріїв. Алгоритм реалізований як окрема комп'ютерна програма на мові Python. Розроблено та верифіковано метод розрахунку розкиду нечіткої залежної змінної навколо її медіанного значення та знаходження верхньої та нижньої меж нечіткого регресійного рівняння. Доведено, що запропоновані методи можуть виступати альтернативою найвідомішим методам побудови лінійної регресії, які постулюють нормальний розподіл похибок.

Ключові слова: регресійний аналіз, метод найменших квадратів, метод найменших модулів, теорія нечіткості.

Вступ

На сучасному етапі розвитку аналітичної хімії в блоці методів аналізу все більшого значення набувають інструментальні методи, а в загальнотеоретичному блоці – засоби забезпечення якості вимірювань хімічного складу. Інструментальні методи надають аналітику великі числові масиви, які необхідно зберігати, порівнювати з наявними в базах даних аналогами, обробляти, спираючись на методи інформатики та теорії аналізу даних. Забезпечення якості вимірювань висуває на авансцену аналітичної хімії метрологічну проблематику, у зв'язку з чим особливо актуальними стають проблеми розширення обсягу, підвищення точності та достовірності інформації, що вилучається з результатів вимірювань. Очевидно, що при розв'язанні цього завдання також не обійтися без інтенсивного застосування комп'ютерно-орієнтованих математичних методів. Не буде великим перебільшенням сказати, що наприкінці ХХ – початку ХХІ століття лейтмотивом застосування різних аналітичних методів стало перетворення масивів результатів вимірювань у аналітичні висновки із застосуванням теорії аналізу даних та інформаційних технологій. В умовах, коли об'єкти хімічного аналізу дуже відрізняються, а теоретичні основи використовуваних методів часто належать різним науковим дисциплінам, впроваджені в аналітичну хімію інформаційні технології та методи теорії аналізу даних залишаються, поряд з теорією пробовідбору та пробопідготовки, чи не єдиними елементами, які об'єднують розрізнені розділи аналітичної хімії.

Завдання оцінювання параметрів лінійних регресій залишається однією з найважливіших проблем сучасної статистики та хемоінформатики (хеометрії). Це пов'язано з тим, що вихідні дані, які отримуються в хімічному експерименті, часто мають значний розкид, вимірювання величин предикторів і залежної змінної не рівноточні, зустрічаються викиди (промахи) а також, доволі часто, відсутні теоретичні міркування щодо виду функціональної залежності. Особливу проблему становить також дослідження вибірок з недостатньою кількістю спостережень, а також ситуації, коли важко зробити висновки про закон розподілу похибок. Все це веде до необхідності розгляду низки альтернативних підходів до побудови рівнянь регресій.

Представлена стаття є складовою частиною циклу робіт, в яких описані різні підходи до побудови (полі)лінійних регресійних моделей, зокрема моделей QSAR (Quantitative Structure-Activity Relationship), [1-5].

Слід зазначити, що серед великої кількості підходів, останнім часом, у зв'язку з активізацією робіт в області машинного навчання, зокрема методів на основі нейронних мереж, значна увага приділяється підходам, які засновані на теорії «нечіткості» (fuzzy logic, fuzzy set) [6,7]. На відміну від традиційних теоретико-ймовірнісних підходів, в «нечітких» теоріях розглядається так звана теорія можливостей [8]. При цьому певний розкид вихідних даних вважається закладеним в самій основі розрахункової схеми. У ряді робіт реалізовано багато типових статистичних задач оцінювання та дискримінації на основі теорії нечітких даних. Значну увагу приділено і регресійним моделям. В представленій роботі розглядається альтернативний варіант побудови зваженої нечіткої задачі лінійної регресії, що орієнтована на робастне оцінювання даних.

Лінійні регресійні моделі

Спочатку розглянемо систему лінійних рівнянь вигляду

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_m x_m \quad (1)$$

де $\beta_0, \beta_1, \beta_2, \dots$ – регресійні коефіцієнти, які підлягають оцінюванню. У матричному формулюванні задача (1) має вигляд:

$$Y = X\beta \quad (2)$$

де X – матриця значень аргументів – x_i , Y – вектор-стовпчик відгуків експериментальної системи з матричними елементами y_i , а β – вектор-стовпчик оцінюваних параметрів:

$$X = \begin{pmatrix} 1 & x_{11} & x_{12} & \dots & x_{1m} \\ 1 & x_{21} & x_{22} & \dots & x_{2m} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{N1} & x_{N2} & \dots & x_{Nm} \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_N \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \dots \\ \beta_m \end{pmatrix}, \quad (3)$$

N – обсяг експериментальної вибірки. Оцінка параметрів такої регресії має відому матричну форму методу найменших квадратів (Ordinary Least Squares, **OLS**):

$$\beta_{OLS} = (X^+ X)^{-1} X^+ Y \quad (4)$$

де символ «+» відповідає операції транспонування матриці. Традиційно нерівноточність вимірювань компенсується застосуванням зваженого методу найменших квадратів (Weighted Least Squares, **WLS**) шляхом призначення статистичних ваг, що означає внесення додаткової апріорної інформації (наприклад, про закон розподілу похибок) в досліджувану систему. В цьому випадку система (2) модифікується в такий спосіб:

$$WY = WX\beta \quad (5)$$

де W – квадратна діагональна матриця:

$$W = \begin{pmatrix} w_1 & 0 & \dots & 0 \\ 0 & w_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & w_N \end{pmatrix} \quad (6)$$

Розв'язок рівняння (5) має вигляд:

$$\beta_{WLS} = (X^+ W X)^{-1} X^+ W Y \quad (7)$$

Альтернативним підходом може виступати метод найменших модулів (Least Absolute Deviation, **LAD**) [9], який є завідомо робастним. Основою методу **LAD** є мінімізація функції «модуль» (U) за параметрами регресії

$$U_{LAD}(\beta) = \|Y - X\beta\|_1 = \sum_{i=1}^N |y_i - \beta_0 - \beta_1 x_{i1} - \dots - \beta_m x_{im}| \quad (8)$$

Легко показати, що (8) можна представити як зважений за допомогою функції «модуль» метод **WLS**:

$$U_{LAD}(\beta) = \|Y - X\beta\|_1 = (Y^+ - \beta^+ X^+)W(Y - X\beta) \quad (9)$$

При цьому ваговий фактор має наступне діагональне представлення:

$$W = \text{diag}(1/|y_i - \beta_0 - \beta_1 x_1 - \dots - \beta_m x_m|), i = 1, \dots, N \quad (10)$$

В рамках методу варіаційно-зважених квадратичних наближень [9] задача мінімізації (8) розв'язується ітераційно.

Теорія нечітких множин, як і відповідна арифметика, в простішому випадку, ґрунтується на «трикутному» нечіткому поданні числа \tilde{Z} у вигляді трійки чисел $\tilde{Z} \equiv (Z, a, b)$. Тут Z – мода або чітке значення трикутного числа, a – ліва, а b – права границі числа \tilde{Z} .

$$a \leq Z \leq b \quad (11)$$

Такі величини іноді називають LR-числом (лівим-правим числом). При цьому вводиться так звана функція належності, яка дозволяє інтерпретувати будь-яке дійсне число (скажімо X) по відношенню до числа \tilde{Z} :

$$\mu_{\tilde{Z}}(X) = \begin{cases} \frac{X - a}{Z - a}, & X \in [a, Z] \\ \frac{X - b}{Z - b}, & X \in [Z, b] \\ 0, & Z \notin [a, b] \end{cases} \quad (12)$$

Для (12) зазвичай використовується лінгвістичне трактування. Наприклад, якщо $\mu_{\tilde{Z}}(X) \sim 1$, то X – «майже Z ». Якщо X близько до границь $X \sim a$ або $X \sim b$, то X «майже не Z ». У центрі інтервалу (наприклад $X \sim |Z + a|/2$): X чи то Z , чи то не Z .

Як приклад, зобразимо (Рис. 1) нечітке трикутне число $\tilde{Z} \equiv (3, 0, 8)$.

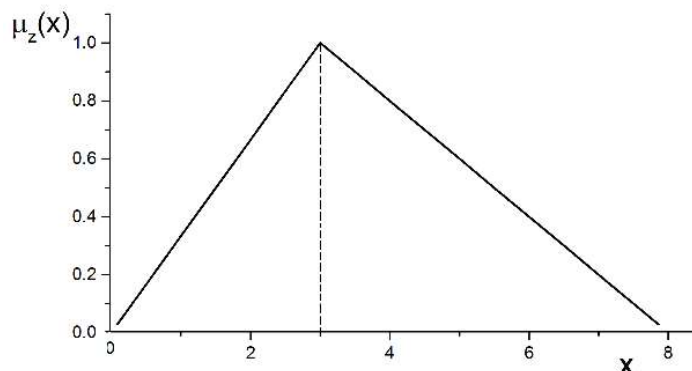


Рис. 1. Графік функції належності для числа $\tilde{Z} \equiv (3, 0, 8)$

Fig. 1. Graph of the membership function for the $\tilde{Z} \equiv (3, 0, 8)$ number

Арифметика, що реалізує операції з нечіткими числами, описана в ряді робіт (див. наприклад [10]).

Побудова відповідних рівнянь нечіткої регресії на основі методу найменших квадратів (Fuzzy Least Squares, **FLS**) вперше було описано в рамках лінійного програмування [11]. Пізніше, в роботі [12], було запропоновано розв'язок **FLS** в термінах класичного методу найменших квадратів.

У самій постановці завдання нечіткої регресії можна виявити кілька варіантів (Табл. 1). Загальний огляд різних підходів наведено, наприклад, в [13,14].

Таблиця 1. Основні підходи щодо реалізації нечіткої регресії
Table 1. Basic approaches to implementing fuzzy regression

	Залежна змінна, Y	Незалежні змінні, X	Коефіцієнти регресії, β
1	нечітка,	нечіткі,	нечіткі,
2	нечітка,	нечіткі,	чіткі,
3	нечітка,	чіткі,	нечіткі,

У представленій роботі ми розглядаємо третій варіант реалізації теорії як найбільш типовий у хімічній науці. Тобто, предиктори X є чіткими, а відгук – нечітким LR-числом $\tilde{Y} \equiv (Y, \lambda, \eta)$. Очевидно, що при цьому параметри регресії також повинні бути нечіткими. Таким чином, апроксимація, за аналогією (2), зв'язується з рішенням лінійної задачі

$$\tilde{Y} = X\tilde{\beta} \quad (13)$$

Для знаходження розв'язків за допомогою методу **OLS** (наприклад [12,15]) використовується узагальнене представлення функції, що мінімізується

$$D^2 = \|X\beta - Y\|_2^2 + \|Xb - \lambda\|_2^2 + \|Xc - \eta\|_2^2 \quad (14)$$

Тут $\|\dots\|_2^2$ – квадрат евклідової норми вектора. Цілком реалістично припустити, що відгук представляє собою симетричне нечітке число $\tilde{Y} \equiv (Y, (Y - \lambda), (Y + \lambda))$, і, як наслідок, коефіцієнти регресії також є симетричними нечіткими числами, $\tilde{\beta} \equiv (\beta, (\beta - b), (\beta + b))$. Відповідна (подібна **OLS**) система рівнянь **FLS**:

$$\frac{\partial D^2}{\partial \beta^+} = 0, \quad \frac{\partial D^2}{\partial b^+} = 0 \quad (15)$$

дозволяє знайти шукані величини аналогічно (4).

Треба відзначити, що хоч певний розкид даних і закладено у **FLS** від самого початку, тим не менш, ця теорія, строго кажучи, не є робастною, оскільки спирається на стандартний метод найменших квадратів. У даній роботі розглянуто зважений варіант **FLS** (weighted **FLS**, **WFLS**). В цьому випадку узагальнений вираз (14) на основі (12), для симетричних нечітких даних, може виглядати наступним чином:

$$\begin{aligned} D_w^2 = & (\beta^+ X^+ - Y^+) W(X\beta - Y) + \\ & + ((\beta - b)^+ X^+ - (Y - \lambda)^+) W(X(\beta - b) - (Y - \lambda)) + \\ & + ((\beta + b)^+ X^+ - (Y + \lambda)^+) W(X(\beta + b) - (Y + \lambda)) \end{aligned} \quad (16)$$

Очевидно, що за аналогією з (9), виходячи з (14), можна реалізувати також і нечіткий метод найменших модулів (Fuzzy Least Absolute Deviation, **FLAD**). В цьому випадку вагова функція має вигляд (10).

Реалізація методу **WFLS** для функції однієї змінної

У представленій роботі діагональні елементи матриці функцій належності, що характеризують нечіткість даних, а також діагональні елементи матриці ваг методу **WFLS** розраховуються виходячи з наступних міркувань. Припускаючи належність у і х до одного регресійного рівняння, при наявності лінійної функціональної залежності між цими змінними, нахил прямої повинен бути «незмінним» при переході від точки до точки. Таким чином, величини похибок вимірювань можна охарактеризувати як коливання g_{ij} навколо середнього значення. Для оцінювання цих коливань ми розраховуємо нахили прямих, що з'єднують кожен пару точок.

$$g_{ij} = \frac{y_i - y_j}{x_i - x_j}; \quad i = 1, 2, \dots, N; \quad j = 1, 2, \dots, N, \quad i \neq j \quad (17)$$

Таким чином, для кожної точки формується вибірка з $N-1$ нахилів. Вибірki коригуються видаленням нахилів, які найбільше відрізняються від інших. Таке коригування введено для того щоб ідентифікувати такі пари точок-викидів, які вносять похибку у визначенні ваги точки. Для кожної точки знаходиться середнє значення (\bar{g}_i) та стандартне відхилення (σ_{g_i}) відповідної вибірки нахилів. Також знаходиться середнє значення (\bar{g}) та стандартне відхилення ($\sigma_{\bar{g}}$) середніх значень нахилів. В ідеальному випадку, коли усі точки лежать на лінії, середні значення нахилів однакові для усіх точок, а стандартні відхилення усіх вибірок дорівнюють нулю. Отже, найбільші значення стандартних відхилень та найбільші відхилення від середнього відповідають точкам-викидам.

Вагові коефіцієнти (w_i) визначаються на основі стандартних відхилень нахилів. Для знаходження вагових коефіцієнтів використовуємо формулу

$$w_i = \frac{1}{\sigma_i \cdot \exp\left(\left(\frac{\bar{g}_i - \bar{g}}{\sigma_{\bar{g}}}\right)^m\right)} \quad (18)$$

У випадку, коли σ_i є надто малою величиною, для зменшення впливу похибок у її визначенні використовуємо адаптовану вагову функцію:

$$w_i = \frac{1}{\left(\sigma_i \cdot \exp\left(\left(\frac{\bar{g}_i - \bar{g}}{\sigma_{\bar{g}}}\right)^m\right)\right)^{1+0.05 \lg(\sigma_i)}} \quad (19)$$

В формулах (18,19) параметр $m = 2, 4$; також визначає вплив відхилення нахилу від середнього на вагову функцію. В представлених нижче прикладах нами обрано $m = 4$.

Отримані ваги нормуються ($\sum_i w_i = 1$), а знайдені значення ваг використовуємо для визначення функції належності μ (12). При цьому припускаємо, що чим більше значення ваги точки, тим точніше вона визначається, і відповідна область її нечіткості є вужчою.

Визначення виду функції нечіткості зазвичай є досить вільним. У нашій роботі ми розглядали наступну функцію μ як величину, що залежить від ваги точки та різниці між експериментальним і теоретичним значенням залежної величини (Δy):

$$\mu_i = \frac{|\Delta y|}{2} + \frac{a}{bNw_i + 1} \sigma \quad (20)$$

В такому разі вибір параметрів a та b визначає поведінку μ у граничних випадках. Також граничні значення, за нашими представленнями, мають бути орієнтовані на стандартне відхилення регресії σ . Так, з формули (20) випливає, що при виборі параметрів $a = 3$ та $b = 5$, для рівноточного випадку маємо $Nw_i = 1$ і функція належності $\mu_i = \frac{\Delta y + \sigma}{2}$, а у випадку, коли

$$w_i \rightarrow 0 \text{ – максимальна } \mu_i \rightarrow \frac{\Delta y}{2} + 3\sigma.$$

Отримані матриці μ та W , відповідно до рівняння (16), використовуємо для знаходження коефіцієнтів регресії. Таким чином отримуємо три ряди залежних змінних: Y , $(Y + \mu)$, і $(Y - \mu)$, і один ряд незалежних змінних X . За цими вибірками, використовуючи метод **WLS**, згідно (6) і (7) будуємо три регресійних рівняння **FWLS**:

$$\begin{aligned} y &= b_0 + xb_1 \\ (y + \mu) &= b_0^{max} + xb_1^{max} \\ (y - \mu) &= b_0^{min} + xb_1^{min} \end{aligned} \quad (21)$$

Перше рівняння є медіанним рівнянням регресії, а лінії, задані двома іншими рівняннями, обмежують смугу нечіткості лінійної моделі. Для врахування точок-викидів при побудові рівнянь-границь вагові коефіцієнти зводяться до ступеню 0.5 у порівнянні з (18) та (19). Функція належності для точок до регресії є максимальною для точок медіани, а по мірі віддалення від медіани – спадає.

Описані алгоритми реалізовано нами на скриптовій комп'ютерній мові Python.

Тестові розрахунки

Результати тестових розрахунків представлено для ряду модельних даних, що включають певні викиди. В якості характеристик регресійних моделей, крім стандартних відхилень (standard deviation, SD), нами розраховано коефіцієнти детермінації R^2 що характеризують розраховані значення залежної змінної і аналогічні величини отримані за процедурою «Leave-One-Out» LOO – Q^2 [16,17].

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2}, \quad Q^2 = 1 - \frac{\sum_i (y_i - \hat{y}_{i/i})^2}{\sum_i (y_i - \bar{y})^2} \quad (22)$$

В формулах (22) y_i – задані (вхідні) значення залежної змінної, \hat{y}_i – розраховані за отриманим лінійним рівнянням відповідні величини, $\hat{y}_{i/i}$ – величини розраховані за процедурою LOO, а \bar{y} – середнє значення для вибірки $\{y_i\}$. Для методів **WLS** та **LAD** нами розраховано також зважений коефіцієнт кореляції

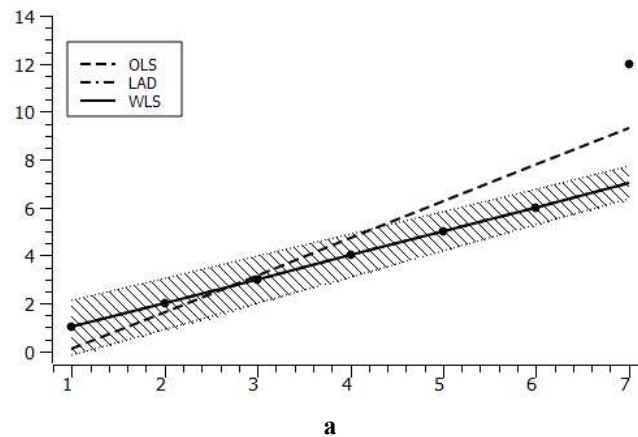
$$R_w^2 = 1 - \frac{\sum_i w_i (y_i - \hat{y}_i)^2}{\sum_i w_i (y_i - \bar{y}_w)^2}, \quad \bar{y}_w = \frac{\sum_i w_i y_i}{\sum_i w_i} \quad (23)$$

де w_i – ваги що були отримані в результаті розрахунку методом **WLS**. В методі **LAD** в якості ваг ми будемо використовувати величини

$$w_i = 1 / |y_i - a_0 - a_1 x_i| \quad (24)$$

При цьому, в методі **LAD**, в формулі (23), при розрахунках суми, будемо пропускати ті точки, через які проходить лінія регресії.

Проілюструємо чисельні дані. Спочатку розглянемо модельні ситуації з явними викидами (рис. 2).



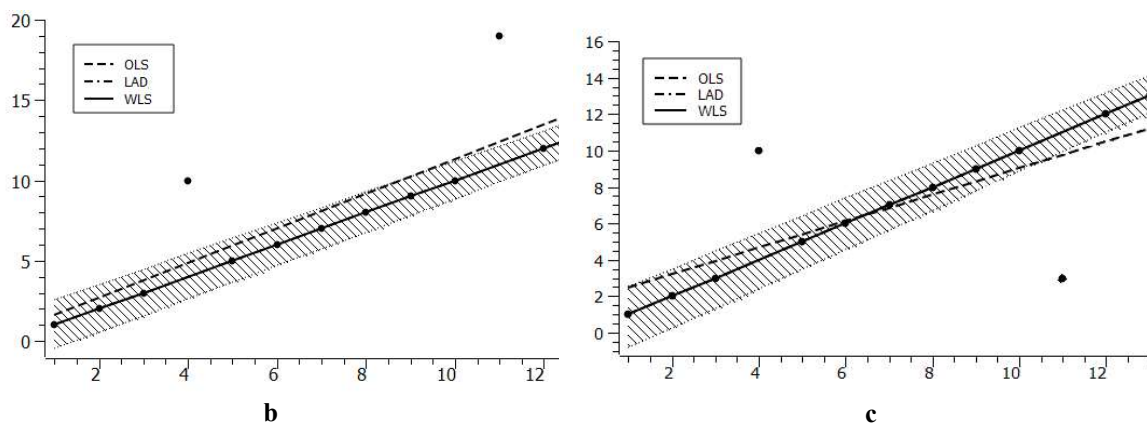


Рис. 2. Графіки-смуги регресійних моделей для вибірок з одним (а) та двома (b, c) викидами. В усіх трьох графіках лінії WLS та LAD співпадають.

Fig. 2. Bar graphs of regression models for samples with one (a) and two (b, c) outliers. In all three graphs, the WLS and LAD lines coincide.

Тут представлено результат апроксимації даних, коли одна точка (рис. 2-а), та дві точки (рис. 2-б, 2-с) суттєво віддалені від основної маси. При цьому основна маса точок точно лежить на лінії. Звісно, що не робастний метод **OLS** в такій ситуації дає рівняння, які значно «розгорнуті» в сторону викидів. Цікаво, що при наявності двох «випадаючих» точок з однієї сторони (рис. 2-б), **OLS** хоч помітно зсуває лінію регресії відносно основної маси точок, однак в цілому дає не такий поганий результат як у випадку з однією «випадаючою» точкою (рис. 2-а). Відзначимо, що результати отримані методом **WLS** близькі до результатів методу **LAD** і на представлених графіках ці лінії співпадають. Два викиди по різні сторони від основної густини точок (рис. 2-с) є особливо складною проблемою для методу **OLS**, тоді як **WLS** та **LAD** знову дають близькі (і найкращі за даної ситуації) результати.

Зафарбована площина навколо медіанної лінії функції відповідає множині точок, що належать до області нечіткості регресійної моделі (відповідає моделі **FWLS**). Можна бачити, що смуга нечіткості досить вузька і мало потерпає від точки викиду.

Отже, у випадку, коли у вибірці наявні один або більше викидів метод **WLS** (як і **LAD**), на відміну від методу **OLS**, зберігає свою робастність.

Типовий приклад регресійних залежностей із розкидом в усіх точках представлено на рис. 3. Результати розрахунків наведено в табл. 2 (для методу **WLS** наведено медіанне рівняння).

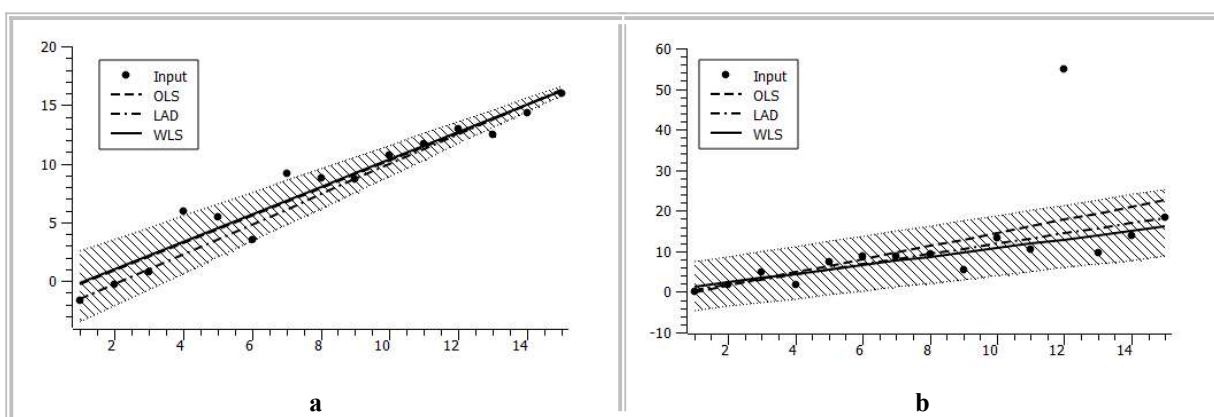


Рис. 3. Графік-смуга регресійних моделей із розкидом (а) без викиду та (б) з викидом
 Fig. 3. Scatter plot of regression models (a) without outliers and (b) with outliers

Таблиця 2. Основні характеристики регресійних моделей для даних, наведених на рис. 3.
Table 2. The main characteristics of regression models for the data shown in fig. 3

Рис. 3-а	Рівняння $y = \beta_0 + \beta_1 x$	R^2	Q^2	$R^2 - Q^2$	R_w^2	SD
OLS	$y = -1.517 + 1.178 \cdot x$	0.9349	0.9154	0.02	–	1.39
LAD	$y = -2.857 + 1.279 \cdot x$	0.9177	0.8618	0.06	0.9896	1.56
WLS	$y = -1.327 + 1.167 \cdot x$	0.9345	0.8812	0.05	0.9630	1.40
Рис. 3-б						
OLS	$y = -1.675 + 1.619 \cdot x$	0.3072	0.0899	0.2	–	10.9
LAD	$y = -0.774 + 1.267 \cdot x$	0.2695	0.1872	0.08	0.9040	11.2
WLS	$y = 0.2567 + 1.054 \cdot x$	0.2277	0.2319	-0.004	0.8587	11.5

На рис.3-а представлено модельні дані із розкидом, але без вираженого викиду. Тут можна бачити, що кожен метод поставляє «своє власне» рівняння, відмінне від інших. Якість рівняння, яка визначається коефіцієнтами детермінації (R^2 , Q^2) та SD приблизно однакова для усіх трьох регресійних моделей. Втім різниця $R^2 - Q^2$, яка характеризує прогностичну здатність рівнянь, трохи гірша для **LAD** та **WLS** ніж для **OLS**. Смуга **FWLS** натурально звужена в області із меншим розкидом.

На рис 3-б представлено дані в яких разом із розкидом присутній також певний (досить значний) викид. В такій ситуації коректні оцінки надано лише методами **LAD** та **WLS**. Звісно, що звичайний коефіцієнт детермінації R^2 не може бути використаний для характеристики отриманих рівнянь. Але запропонована величина R_w^2 , яка враховує ваги точок, характеризує отримані рівняння як задовільні. Відзначимо втім, що рівняння **LAD** та **WLS** суттєво відрізняються вільним членом β_0 .

В останньому прикладі використано класичні дані відомого своїми піонерськими роботами в області робастної статистики П. Х'юбера [18]. Аналіз цієї задачі див. в [19]. Вибірка Х'юбера включає шість точок які відповідають лінійній моделі

$$y = -2 - x \tag{25}$$

В перші п'ять точок внесено випадкову похибку із нульовим середнім значенням і стандартним відхиленням $SD = 0.6$. В шосту точку внесено значно більшу похибку.

Результати розрахунків представлено в табл. 3 та на рис. 4.

Таблиця 3. Характеристики регресійних моделей для задачі Х'юбера.
Table 3. Characteristics of regression models for the Huber problem.

	Рівняння $y = \beta_0 + \beta_1 x$	R^2	Q^2	R_w^2	SD
OLS	$y = 0.068 - 0.081 \cdot x$	0.0824	-13.42	–	1.39
LAD	$y = -0.0333 + 0.0033 \cdot x$	-0.0128	-14.04	-0.0067	1.46
WLS	$y = -1.325 - 0.608 \cdot x$	-4.470	-9.98	0.6610	3.38

Не дивно, що метод **OLS** показав неадекватні результати. Можна бачити, що навіть **LAD**, який зазвичай є робастним виявився неадекватним. Вочевидь це пов'язано із малою кількістю точок в лінійній частині даних і результуюча лінія **LAD** розгорнута до точки-викиду як і **OLS** (див. рис. 4). Найкращі, хоч і не «блискучі» ($R_w^2 = 0.6610$) результати виявив метод **WLS**. Отримане **WLS** регресійне рівняння якісно близьке до незбуреного первинного рівняння (25). В цьому прикладі точка, що відповідає викиду веде до значного розширення смуги невизначеності **FWLS**.

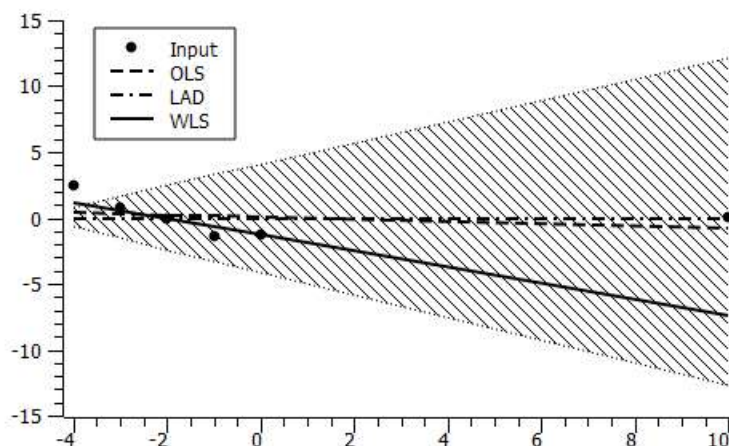


Рис. 4. Лінійні регресійні моделі для прикладу Х'юбера
 Fig. 4. Linear regression models for Huber's example

Висновки

У представленій роботі запропоновано підхід для знаходження вагових параметрів зваженого методу найменших квадратів, а також максимальних і мінімальних можливих параметрів регресійного рівняння. Для призначення вагових коефіцієнтів використовується принцип рівності нахилів прямої лінії між кожною парою точок. Цей метод досить адекватно відсіює точки-викиди, призначаючи їм найменші значення ваг. Відповідно, методи **WLS** у розглянутих в даній роботі прикладах також показали доволі високу робастність. Їхньою перевагою, на відміну від методу найменших модулів, є неперервність та диференційованість функціоналу, який мінімізується. На підставі знайдених ваг запропоновано метод розрахунку розкиду нечіткої вимірюваної величини навколо її середнього значення та знаходження верхньої та нижньої меж нечіткого регресійного рівняння.

Зважаючи на вищезазначені переваги, ці методи можуть виступати альтернативою найвідомішим методам побудови регресії, які постулюють закони розподілу похибок Гауса або Лапласа. Тому їх можна запропонувати до застосування у випадках, коли ідентифікація похибки у вимірюваннях ускладнена та у разі складностей з припущенням про закон розподілу похибок.

Подяки

Робота виконувалась за часткової фінансової підтримки на виконання завдань перспективного плану розвитку наукового напрямку "Математичні науки та природничі науки" Харківського національного університету імені В. Н. Каразіна, № держреєстрації 0121U112886.

Список літератури

1. Onizhuk M.O., Ivanov V.V., Panteleimonov A.V., Kholin Yu.V. Alternative Methods for Constructing of Linear Regressions. *Method and object Chemical Analysis*. 2017, 12(3), 105-111. <https://doi.org/10.17721/moca.2017.105-111>
2. Berdnyk M. I., Onizhuk M. O., Ivanov V. V. Methods for building linear regression equations in the "structure-property" problems. *Kharkov University Bulletin. Chemical Series*. 2018, 30 (53), 6-17. <https://doi.org/10.26565/2220-637X-2018-30-01>
3. Berdnyk M.I., Zakharov A.B., Ivanov V.V. Application of L₁-regularization approach in QSAR problem. Linear regression and artificial neural networks. *Method and Object Chemical Analysis*. 2019, 14(2), 79-90. <https://doi.org/10.17721/moca.2019.79-90>
4. Zakharov A.B., Dyachenko A.V., Ivanov V.V. Topological Characteristics of Iterated Line Graphs in QSAR Problem: Octane Numbers of Saturated Hydrocarbons. *Journal of Chemometrics*. 2019, 33 (9), e3169. <https://doi.org/10.1002/cem.3169>

5. Zakharov A. B., Tsarenko D. K., Ivanov V. V. Topological characteristics of iterated line graphs in the QSAR problem: a multigraph in the description of properties of unsaturated hydrocarbons. *Struct Chem.* 2021, 32, 1629-1639. <https://doi.org/10.1002/cem.3169>
6. Fuzzy logic in Chemistry. Rouvray D.H. ed. Academic Press, London, 1997, 364 p.
7. Zadeh L. A. The Concept of a Linguistic Variable and its Application to Approximate Reasoning-I. *Information Sciences.* 1975, 8, 199-249. [https://doi.org/10.1016/0020-0255\(75\)90036-5](https://doi.org/10.1016/0020-0255(75)90036-5)
8. Dubois D., Prade H. Possibility Theory, Probability Theory and Multiple-Valued Logics: A Clarification. *Annals of Mathematics and Artificial Intelligence.* 2001, 32, 35–66. <https://doi.org/10.1023/A:1016740830286>
9. Bloomfield P., Steiger W.L. Least Absolute Deviations. Theory, Applications and Algorithms. Boston: Birkhäuser, 1983, 351 p.
10. Hanss M. Applied Fuzzy Arithmetic. An Introduction with Engineering Applications. Springer-Verlag Berlin Heidelberg 2005, 256 p.
11. Tanaka H., Uegima S., Asai K. Linear Regression Analysis with Fuzzy Model. *IEEE Trans. on Systems, Man and Cybernetics.* 1982,12, 903–907. <http://dx.doi.org/10.1109/TSMC.1982.4308925>
12. Diamond P. Fuzzy Least Squares. *Information Sciences.* 1988, 46, 141–157. [https://doi.org/10.1016/0020-0255\(88\)90047-3](https://doi.org/10.1016/0020-0255(88)90047-3)
13. de Andrés-Sánchez J. Fuzzy Regression Analysis: An Actuarial Perspective in Fuzzy Statistical Decision-Making Theory and Applications. Springer. 2016. 173-201. https://doi.org/10.1007/978-3-319-39014-7_11
14. Conventional and fuzzy regression theory and engineering applications. Hrissanthou V., Spiliotis M. (eds). Nova Science Publishers Inc, New York. 2018, 332 p.
15. Haggag M.M.M. A New Fuzzy Regression Model by Mixing Fuzzy and Crisp Inputs. *American Review of Mathematics and Statistics.* 2018, 6(2), 9-25. <https://doi.org/10.15640/arms.v6n2a2>
16. Golbraikh A., Tropsha A. Beware of Q2! *Journal of Molecular Graphics and Modelling.* 2002, 20, 269-276. [https://doi.org/10.1016/S1093-3263\(01\)00123-1](https://doi.org/10.1016/S1093-3263(01)00123-1)
17. Alexander D.L.J., Tropsha A., Winkler D.A. Beware of R2: Simple, Unambiguous Assessment of the Prediction Accuracy of QSAR and QSPR Models. *J. Chem. Inf. Model.* 2015, 55 (7), 1316–1322. <https://doi.org/10.1021/acs.jcim.5b00206>
18. Huber P.J. Robust Statistics. J. Wiley and sons, New York. 1981, 308 p.
19. Kholin Yu. V. A quantitative physicochemical analysis of complexation in solutions and on the surface of complexing silicas: meaningful models, mathematical methods and their application. Kharkiv, Folio, 2000, 294 p. (in Rus).

Надійшло до редакції 27.06.2022

Прийнято до друку 12.09.2022

A. V. Panteleimonov, D. O. Anokhin, V. V. Ivanov. Robust evaluation of regression parameters. The fuzzy theory and other models.

V. N. Karazin Kharkiv National University, 4 Svobody sq., Kharkiv, 61022, Ukraine

Linear regression parameters based on fuzzy theory are compared with other statistical approaches. A new algorithm of a simple weighted least squares method, independent of a priori information, is proposed. The algorithm was verified on model data, and its adequacy was confirmed with the use of standard criteria. The algorithm has been implemented as Python language computer program. New method of calculation of the scatter of fuzzy dependent variable around its median value, as well as the upper and lower bounds of fuzzy regression equations have been developed and verified. Proposed methods are shown to be useful alternatives to the most popular methods for constructing linear regression, which assume a normal distribution of errors.

Keywords: regression analysis, Least Square method, Least Absolute Deviation method, fuzzy theory.

Kharkiv University Bulletin. Chemical Series. Issue 38 (61), 2022

EXAMINATION OF MIXED SODIUM DODECYL SULFATE – 1-BUTANOL MICELLES WITH MOLECULAR DYNAMICS**I. V. Kononenko^a, V. S. Farafonov^b**

V. N. Karazin Kharkiv National University, 4 Svobody sq., Kharkiv, 61022, Ukraine

a) ✉ kononenko2109@gmail.com <https://orcid.org/0009-0009-7762-6005>b) ✉ farafonov@karazin.ua <https://orcid.org/0000-0003-0785-9582>

Properties of mixed micelles composed of sodium *n*-dodecyl sulfate surfactant and 1-butanol co-surfactant were examined using molecular dynamics simulation for three experimentally observed compositions. The aggregation number varied in the range of 30 to 64, and the co-surfactant content was 75 to 32 molecules, respectively. The composition of the bulk water – 1-butanol mixture matched the experimentally studied concentrations. All-atom potential models and explicit solvent were used in order to capture the most structural details. Such characteristics as micelle size, co-surfactant location, depth of water penetration were determined. The simulated micelles were found to have compact ellipsoidal shape, except of the most co-surfactant-rich ones, which were elongated. The equilibrium composition differed little from the initial one indicating consistence of the simulations. The co-surfactant molecules were observed in different positions: either situated on the surface of surfactant aggregate or having hydrocarbon tails immersed in it. Still, the hydroxyl groups stayed on the micelle surface and preserved contact with bulk solution. Water molecules were found penetrating the mixed micelles up to 0.3 – 1.0 nm from their center. Degree of counterion binding was estimated and its dependence on micelle composition was derived. It was found to decrease with the co-surfactant content in a non-linear manner. Specifically, for co-surfactant content up to 34% the counterion binding was almost similar to that of pure sodium *n*-dodecyl sulfate micelles; the value gradually fade at higher 1-butanol content. Several positions of the boundary between bound and free counterions were considered. The approach proposed here may be employed for examination of other mixed micelles, as well.

Keywords: radial distribution function, micelle size, hydration, degree of counterion binding.

Introduction

At present, surfactant micellar solutions found versatile application. They serve as cleaning agents, disinfection aids, emulgators, as reaction media, and so forth. One of their valuable features is a variety of characteristics, which can be tuned to achieve desired properties: surfactant head group and hydrocarbon radical, counterion, concentration, ionic strength. Introduction of a co-surfactant provides a new degree of freedom, which extends the range of available properties of micelles and bulk solution even more.

In this paper we examine the microscopic structure and counterion binding of mixed micelles made of sodium *n*-dodecyl sulfate (SDS) surfactant and 1-butanol (BuOH) co-surfactant (Fig. 1). On one hand, this is a relevant choice because SDS is among the most widely used surfactants in technology and laboratory practice. On the other hand, extensive experimental data was collected for this system at various concentrations of the components. [1–3] It is summarized in a recent review. [4] However, intriguingly, for most cases the available data is fragmentary. Often the micelle size and composition are reported (radius, aggregation number N_{SDS} and number of co-surfactant molecules N_{BuOH}), but there is no information about degree of counterion binding. Oppositely, when counterion binding was examined, there is no data about the aggregation number and co-surfactant content. This shows that our knowledge of properties of this object is still limited.

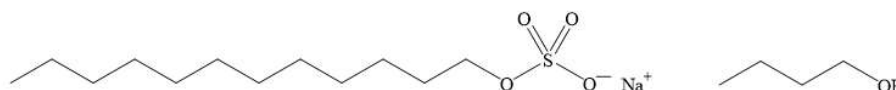


Figure 1. Structure formulas of SDS surfactant and 1-butanol co-surfactant.

Here we will use the method of molecular dynamics (MD) to get a new insight. It is a powerful instrument for identifying microscopic structure of a molecular system. When applied to micelles it can reveal water penetration into the hydrocarbon core, ordering of surfactant molecules, counterion binding, dynamics of self-assembly *etc.* At present, single-component surfactant micelles are thoroughly examined by this method. [5] Computational studies of the mixed micelles are also found in literature, although in much lesser amount. [6–8]

Notably, the chosen system has not yet been studied with MD. The closest match is the MD simulation of mixed micelles made of SDS and medium to higher alcohols (hexanol to hexadecanol) [8]. The authors were focused on structure of the micelles and its dependence on composition. Transitions between different micelle shapes were identified. Importantly, coarse-grained potential models were used there. On one hand, it allowed cover long time scale (microseconds), on the other hand, these models lack detail needed to study ion binding. The mixture of SDS and 1-butanol was investigated, as well, but in an interfacial setup only. [9] There, the effect of the co-surfactant on surface tension and density profile of the surfacial film was identified.

For the current study we chose three compositions of mixed micelles, which were experimentally observed: $N_{\text{SDS}} = 30$ and $N_{\text{BuOH}} = 75$ [1]; $N_{\text{SDS}} = 40$ and $N_{\text{BuOH}} = 40$ [1]; $N_{\text{SDS}} = 64$ and $N_{\text{BuOH}} = 32$ [2]. This choice covers a wide range of the co-surfactant content: from 71% to 33% by amount. For all of these systems the degree of counterion binding was not reported, and micelle radius was not reported for the last one.

Simulation methodology

The simulation cells consisted of a preformed $\text{DS}^- + \text{BuOH}$ aggregate placed in a water – 1-butanol mixture with distributed Na^+ counterion. The initial shape of the aggregate was a rectangular bilayer made of alternating DS^- ions and BuOH molecules. This compact shape can be easily generated and shortens the time period needed to obtain an equilibrated micelle. On one hand, close packing of hydrophobic tails resembles that in a micelle. On the other hand, the initial placement of BuOH molecules within the aggregate allows avoid waiting for their sorption from the bulk solution. The cells were cubic-shaped with edge length 8.5 nm for $N_{\text{SDS}} = 64$ and 8 nm otherwise. One of simulation cells is shown in Fig. 2A. The number of 1-butanol molecules that should be placed in the bulk solution was deduced from the conditions of the experiment, where given composition of mixed micelles was observed. The calculations were done via Eq. 1 and the relevant experimental data is collected in Table 1. The principle is as follows. Let us consider 1 L of the system. Amount of BuOH in the bulk solution equals the difference between its total amount, $c_{\text{BuOH}} \cdot 1 \text{ L}$, and its amount in micelles. The latter equals the amount of SDS in micelles, $(c_{\text{SDS}} - \text{CMC}) \cdot 1 \text{ L}$, multiplied by the ratio $N_{\text{BuOH}} / N_{\text{SDS}}$. Here CMC is critical micelle concentration of SDS; it was taken as 0.004 mol/L because in 1-butanol solutions it decreases compared with pure water [4]. Repeating the calculation for the volume of bulk solution in the MD cell V_{sol} gives the number of BuOH molecules to place in it. Because the micelle occupies a considerable fraction of the cell volume, V_{sol} should be found as the difference between volume of the whole cell V_{cell} and volume of the micelle inside V_{mic} . The latter can be approximated as the volume of a parallelepiped and measured by visual examination, *e.g.*, in VMD software. [13]

$$N_{\text{BuOH}_{\text{sol}}} = N_A \left(c_{\text{BuOH}} - (c_{\text{SDS}} - \text{CMC}) \frac{N_{\text{BuOH}}}{N_{\text{SDS}}} \right) (V_{\text{cell}} - V_{\text{mic}}) \quad (1)$$

where $N_{\text{BuOH}_{\text{sol}}}$ is number of BuOH molecules added in the bulk solution, N_A is Avogadro number.

Table 1. Compositions of the simulated cells and relevant experimental data.

system	N_{SDS}	N_{BuOH}	c_{SDS} , mol/L	c_{BuOH} , mol/L	$V_{\text{cell}} - V_{\text{mic}}$, nm^3	$N_{\text{BuOH}_{\text{sol}}}$	N_{water}
1 [1]	30	75	0.0316	0.764	480	201	15046
2 [1]	40	40	0.033	0.326	480	86	15683
3 [2]	64	32	0.2	0.2	570	35	18936

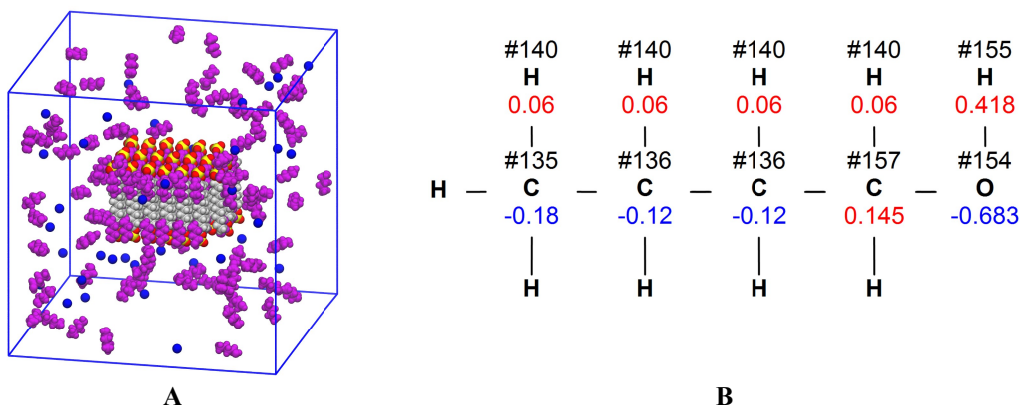


Figure 2. **A:** Simulation cell for micelle with $N_{\text{SDS}} = 40$, $N_{\text{BuOH}} = 40$. Na^+ ions are colored blue, BuOH molecules are colored violet. **B:** Used OPLS-AA potential model of 1-butanol.

Atom–atom interactions in the system were described using OPLS-AA force field. The potential model for SDS was taken from literature [10]. The one for BuOH was composed using the built-in atom types and atomic point charges, which are empirical and optimized for reproducing properties of organic liquids [11]. It is depicted in Fig. 2B. Water was represented by means of the SPC model.

The simulations were carried out at standard conditions ($T = 25^\circ\text{C}$, $p = 1$ bar), which were maintained using Berendsen thermostat and barostat. Electrostatic interactions were calculated with the PME method, while van der Waals interactions were cut at 1 nm. Three-dimensional periodic boundary conditions and bond constraints were imposed. The simulations were 50 ns long, time step equaled 2 fs. GROMACS 5.1 software was used [12].

Results and discussion

As the initial micelle shape was artificial (a bilayer), the beginning of each simulation was devoted to equilibration of the micelle. The distribution of 1-butanol between micelle and solution was equilibrated, as well. To monitor the progress we calculated number of BuOH molecules staying in contact with DS^- ions (that is located within 0.4 nm of any DS^- ion). This number shows the amount of co-surfactant in the micelle. The graphs are shown in Fig. 3.

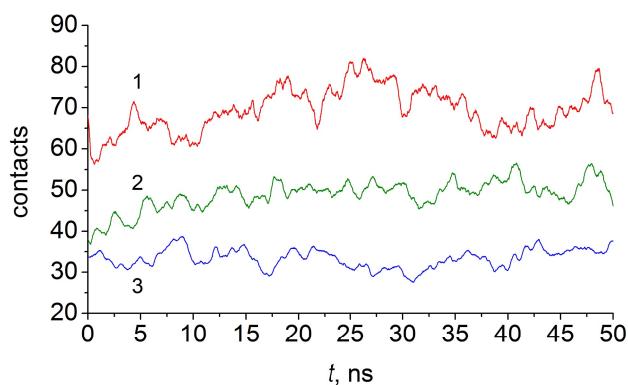


Figure 3. Evolution of number of contacts between DS^- ions and BuOH molecules for micelles of different composition. 1 — $N_{\text{SDS}} = 30$, 2 — $N_{\text{SDS}} = 40$, 3 — $N_{\text{SDS}} = 64$.

We conclude that the original composition of the micelles was mostly preserved. For systems #1, #2, and #3 the average number of BuOH molecules equals 71.4, 50.5, and 33.1, respectively, that is close to the originally placed. This additionally proves the correctness of the computational model. The graphs were also used to identify the end of equilibration phase, which was recognized at 20 ns. The quantities below were computed over the 20–50 ns interval.

The final configurations of two micelles are shown in Fig. 4. Hydroxyl groups of BuOH molecules are located among the surfactant head groups. The butyl radicals can be found either lying on the surface of the SDS aggregate or protruding inside it.

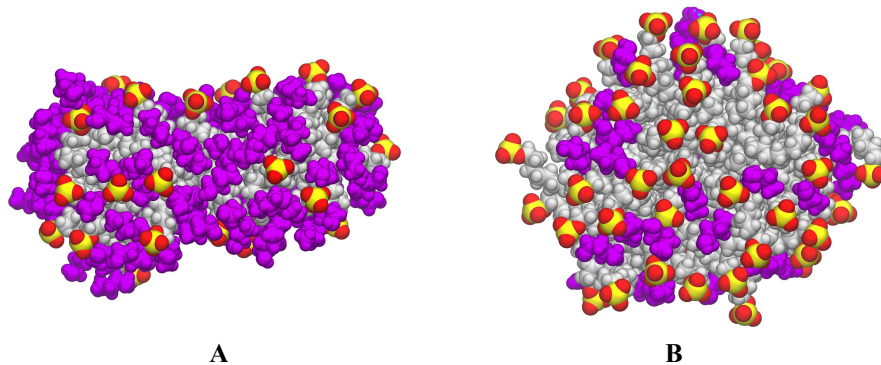


Figure 4. Final configurations for micelle with $N_{\text{SDS}} = 30$, $N_{\text{BuOH}} = 75$ (A) and $N_{\text{SDS}} = 64$, $N_{\text{BuOH}} = 32$ (B). BuOH molecules are colored violet.

To estimate size of the micelles we calculated the radial distribution function (RDF) of S atoms of DS^- anions with respect to micelle center of mass, Fig. 5A. The latter was approximated by the center of mass of DS^- anions, which is justified by uniform distribution of the co-surfactant across the micelle.

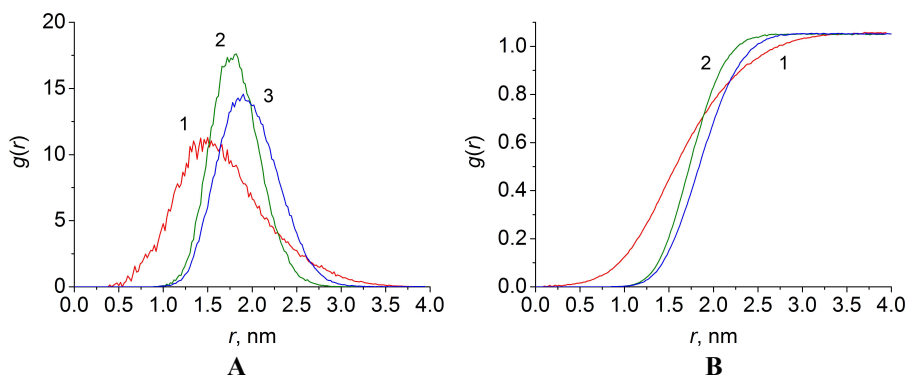


Figure 5. A: RDF of S atoms (A) or water molecules (B) with respect to micelle center of mass. 1 — $N_{\text{SDS}} = 30$, 2 — $N_{\text{SDS}} = 40$, 3 — $N_{\text{SDS}} = 64$.

Abscissa of the peak of the RDF shows the most probable position of head groups and, thus, indicates the radius of micelle. Determined in such a way it equals 1.54 nm, 1.78 nm, 1.88 nm for $N_{\text{SDS}} = 30$, 40, 64, respectively. The first two values agree quite good with the experimental numbers 1.71 nm and 1.68 nm, respectively. The last value (1.88 nm) also equals the radius of pure SDS micelles with $N_{\text{SDS}} = 60$.

The curve of micelle with $N_{\text{SDS}} = 30$ is notably wider than the other. This is because of its pronouncedly prolate shape, as is proven by visual inspection. This elongation is the probable reason why experimental radius is 0.2 nm larger than the one determined from RDF. Importantly, visual inspection also reveals that the SDS content occasionally reversibly splits into two fragments bridged by BuOH molecules for short times. Oppositely, the two other micelles are less fluid and have a more compact, although still ellipsoidal, shape.

More detailed information about co-surfactant distribution in micelles is provided by radial distribution functions of particular 1-butanol atoms with respect to micelle center of mass. The RDF's for hydroxyl O atom and C atoms in positions 2 and 4 are shown in Fig. 6. The peaks of the O atom are located ca. 0.2 nm closer to micelle center than those of S atoms of DS^- , which corresponds to location among ester O atoms of DS^- . The C^2 and C^4 atoms are in average immersed deeper into micelles than hydroxyl O atoms, and this is more pronounced for $N_{\text{SDS}} = 30$: the peak of C^4 atoms is 0.35 nm ($N_{\text{SDS}} = 30$) or 0.25 nm ($N_{\text{SDS}} = 40, 64$) behind that of the O atoms. The numbers are shorter than the $\text{C}^4 - \text{O}$ distance in a stretched BuOH molecule (0.5 nm) showing that in average the molecules are inclined or bent. Considering Fig. 4 we conclude that BuOH molecules can both lie on the surface of DS^- aggregate and be immersed in it, giving average $\text{C}^4 - \text{O}$ distance projection to

micelle radius less than 0.5 nm. Still, in micelles with $N_{\text{SDS}} = 40$ and 64 the co-surfactant does not reach the micelle interior: $g(r) \approx 0$ at $r < 0.6$ nm. The micelles with $N_{\text{SDS}} = 30$ are more labile and rich with co-surfactant, so BuOH can be found in all positions there.

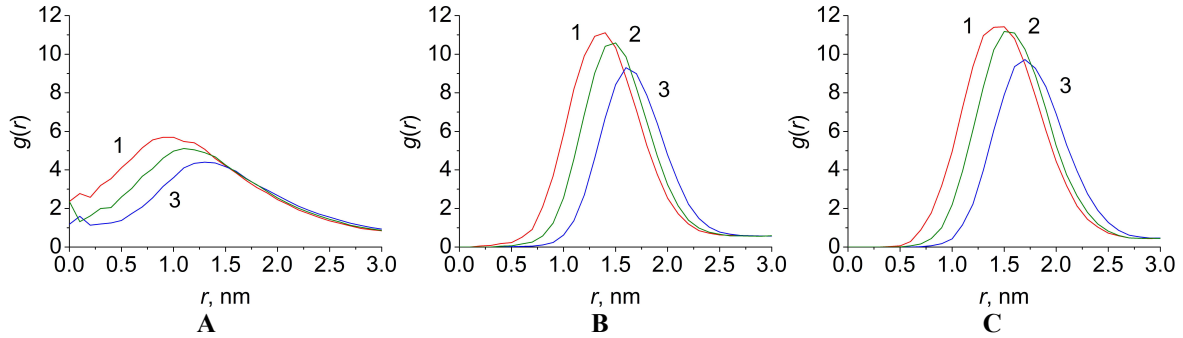


Figure 6. RDF of BuOH C⁴ (1, red), C² (2, green), and O (3, blue) atoms with respect to the micelle center. **A:** $N_{\text{SDS}} = 30$; **B:** $N_{\text{SDS}} = 40$; **C:** $N_{\text{SDS}} = 64$.

The simulations were done with explicit water, therefore, the distribution of water in micelle can be examined. It was described using RDF of water molecules with respect to the micelle center, Fig. 5B. Compared to the bulk solution, the density of water decreases by 2 times at the distance of 1.58 – 1.84 nm from micelle center. This indicates that the head groups and adjacent methylene groups of the micelles are strongly hydrated. Except of $N_{\text{SDS}} = 30$, micelles become almost dry at 1.1 nm (RDF is $< 1\%$ of the bulk value) and completely dry at 0.85 nm. The curve for micelles with $N_{\text{SDS}} = 40$ almost coincides with that of pure SDS micelles having $N_{\text{SDS}} = 60$. Therefore introduction of 1-butanol up to 50% ($N_{\text{SDS}} = N_{\text{BuOH}} = 40$) did not affect the micelle structure considerably. Contrary, low aggregation number and prolate shape of the micelle with $N_{\text{SDS}} = 30$ makes its hydrocarbon core rather thin.

The other question of interest is counterion binding by mixed micelles. To calculate the degree of counterion binding β it is necessary to place the boundary between bound and free Na^+ ions. We used the position of the first and the second minima of the RDF between Na^+ ions and DS^- S atoms as the criterion, which equal 0.44 nm and 0.7 nm, respectively. The graph for the case $N_{\text{SDS}} = 64$ is shown in Fig. 7A; in other cases RDF has similar abscissas of minima and maxima. An even longer distance of 0.9 nm was considered, as well. The ions located closer than the chosen boundary position d were considered as bound and contributed to β . The values are collected in Table 2 and depicted in Fig. 7B. For comparison the data for pure SDS micelles is shown, as well [10]. Molar fraction of BuOH was computed using equilibrium N_{BuOH} values instead of the initial ones.

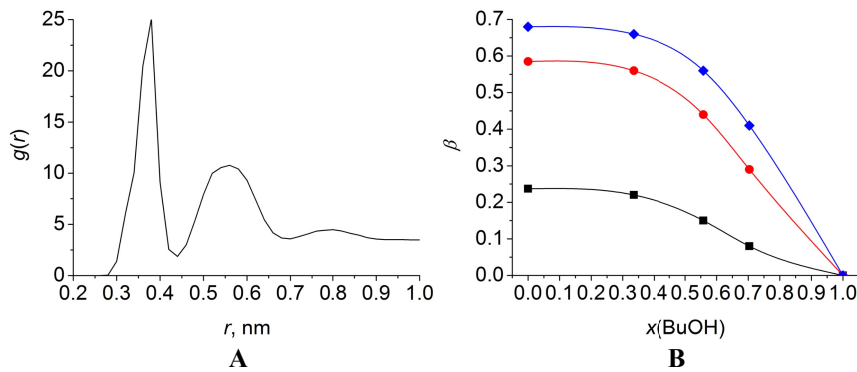


Figure 7. **A:** RDF between S atoms and Na^+ ions. **B:** Dependence of degree of counterion binding on molar fraction of 1-butanol in mixed micelle. Black squares: $\beta(0.44 \text{ nm})$, red circles: $\beta(0.7 \text{ nm})$, blue diamonds: $\beta(0.9 \text{ nm})$. Lines are produced by spline interpolation.

Table 2. Numbers of bound Na^+ ions and corresponding degrees of counterion binding for different positions of the boundary and micelle compositions.

N_{SDS}	x_{BuOH}	$d = 0.44 \text{ nm}$		$d = 0.7 \text{ nm}$		$d = 0.9 \text{ nm}$	
		N	β	N	β	N	β

30	0.70	2.3	0.08	8.7	0.29	12.3	0.41
40	0.56	5.8	0.15	17.5	0.44	22.4	0.56
64	0.34	14.1	0.22	35.6	0.56	42.0	0.66
60	0	14.2	0.24	35.1	0.59	40.7	0.68

A clear monotonic dependence is observed: increase of 1-butanol content lowers counterion binding. However, the effect becomes significant only when the co-surfactant content is considerable: even at $x_{\text{BuOH}} = 0.34$ β remains almost equal to that of pure SDS micelles. We suggest the following explanation. The micelle with $x_{\text{BuOH}} = 0.34$ contains $N_{\text{SDS}} = 64$ and $N_{\text{BuOH}} = 32$. Size and shape of its DS^- aggregate differ little from those of pure SDS micelle with $N_{\text{SDS}} = 60$, leading to similar counterion binding. The additional 1-butanol is distributed over the surface (Fig. 4B). In two other mixed micelles the BuOH molecules incorporate into the hydrocarbon core to larger extent and pull apart the surfactant head groups. As a result, surface charge density decreases, which weakens attraction of Na^+ .

For the considered compositions there was no experimental data about β to compare. However, $\beta = 0.39$ was reported for a system with $c_{\text{SDS}} = 0.02$ mol/L, $c_{\text{BuOH}} = 0.8$ mol/L that is quite close to system #1 (Table 1). [4] For pure micelles the value 0.71 was reported there. Both values agree with matching simulation results for boundary position at 0.9 nm.

Conclusions

Mixed micelles made of sodium *n*-dodecyl sulfate and 1-butanol at standard conditions were simulated by means of molecular dynamics. Three compositions were considered with molar fraction of the co-surfactant 33% to 71%. 1-butanol was found located both on the surface of micelles and incorporated in their hydrocarbon core. Still, its hydroxyl groups stay on the micelle surface among surfactant head groups. Degree of counterion binding was found reducing with increase of 1-butanol content. However, up to molar fraction of 33% the effect was small, and counterion binding was similar to that of the single-component micelles. The used approach may be utilized for examination of other mixed micelles, as well.

Acknowledgements

The authors are grateful to the Ministry of Education and Science of Ukraine for the financial support of this study via grant #0120U101064, and to Dr. Dmitry Nerukh (Aston University, Birmingham, UK) for the generously provided computational resources.

References

1. Almgren, M.; Swarup, S. Size of sodium dodecyl sulfate micelles in the presence of additives i. alcohols and other polar compounds. *J. Colloid Interface Sci.* **1983**, *91*, 256–266. [https://doi.org/10.1016/0021-9797\(83\)90330-2](https://doi.org/10.1016/0021-9797(83)90330-2).
2. Caponetti, E.; Martino, D. C.; Floriano, M. A.; Triolo, R. Localization of *n*-Alcohols and Structural Effects in Aqueous Solutions of Sodium Dodecyl Sulfate. *Langmuir* **1997**, *13*, 3277–3283. <https://doi.org/10.1021/la961013m>.
3. Kharchenko, A. Yu. Composition of the sodium dodecylsulfate – 1-pentanol mixed micelles as determined using acid-base indicators. *Kharkov Univ. Bull., Chem. Ser.* **2016**, *27*(50), 5–15. <https://doi.org/10.26565/2220-637X-2016-27-01>.
4. Kharchenko, A. Yu.; Kamneva, N. N.; Mchedlov-Petrosyan, N. O. The properties and composition of the SDS – 1-butanol mixed micelles as determined via acid-base indicators. *Colloids Surf. A* **2016**, *507*, 243–254. <https://doi.org/10.1016/j.colsurfa.2016.08.004>.
5. Taddese, T.; Anderson, R. L.; Bray, D. J.; Warren, P. B. Recent advances in particle-based simulation of surfactants. *Curr. Opin. Colloid Interface* **2020**, *48*, 137–148. <https://doi.org/10.1016/j.cocis.2020.04.001>.
6. Marrink, S. J.; Mark, A. E. Molecular Dynamics Simulations of Mixed Micelles Modeling Human Bile. *Biochemistry* **2002**, *41*(17), 5375–5382. <https://doi.org/10.1021/bi015613i>.
7. Shang, B. Z.; Wang, Z.; Larson, R. G. Molecular Dynamics Simulation of Interactions between a Sodium Dodecyl Sulfate Micelle and a Poly(Ethylene Oxide) Polymer. *J. Phys. Chem. B* **2008**, *112*(10), 2888–2900. <https://doi.org/10.1021/la00022a026>.

8. Mendez-Bermudez, J. G.; Dominguez, H. Structural changes of a sodium dodecyl sulfate (SDS) micelle induced by alcohol molecules. *J. Mol. Model.* **2015**, 22(33), 1–9. <https://doi.org/10.1007/s00894-015-2904-x>.
9. Parra, J. G.; Aray, Y. B.; Iza, P.; Zarate, X.; Schott, E. Behavior of the SDS/1-butanol and SDS/2-butanol mixtures at the water/*n*-octane interface through molecular dynamics simulations. *Chem. Phys.* **2019**, 523, 138–149. <https://doi.org/10.1016/j.chemphys.2019.04.021>.
10. Farafonov, V. S.; Lebed, A. V. Developing and validating a set of all-atom potential models for sodium dodecyl sulfate. *J. Chem. Theory Comput.* **2017**, 13(6), 2742–2750. <https://doi.org/10.1021/acs.jctc.7b00181>.
11. Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OLPS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996**, 118(15), 11225–11236. <https://doi.org/10.1021/ja9621760>.
12. Abraham, M. J.; Murtola, T.; Schulz, R.; Pall, S.; Smith, J. C.; Hess, B.; Lindahl E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, 1–2, 19–25. <https://doi.org/10.1016/j.softx.2015.06.001>.
13. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, 14, 33–38. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5).

Received 09.08.2022

Accepted 12.09.2022

І.В. Кононенко, В.С. Фарафонов. Дослідження змішаних міцел додецилсульфату натрію – 1-бутанолу методом молекулярної динаміки.

Харківський національний університет імені В.Н. Каразіна, хімічний факультет, майдан Свободи, 4, Харків, 61022, Україна



Властивості змішаних міцел, що складаються з поверхнево-активної речовини (ПАР) *n*-додецилсульфату натрію та ко-ПАР 1-бутанолу досліджені методом моделювання молекулярної динаміки для трьох експериментально спостережуваних складів. Число агрегації перебувало в інтервалі від 30 до 64, а вміст ко-ПАР становив від 75 до 32 молекул відповідно. Склад навколишнього розчину вода – 1-бутанол відповідав експериментально дослідженим концентраціям. Використовувалися повноатомні потенціальні моделі та явний розчинник, щоб відтворити більшість структурних деталей. Визначено такі характеристики, як розмір міцели, розташування ко-ПАР, глибина проникнення води. Змодельовані міцели мали компакту еліпсоїдальну форму, за винятком найбільш багатих на ко-ПАР, які були продовгуватими. Рівноважний склад мало відрізнявся від початкового, що вказує на узгодженість моделювання. Молекули ко-ПАР спостерігалися в різних положеннях: розташовані на поверхні агрегату ПАР або занурені в нього вуглеводневими радикалами. Проте гідроксильні групи перебували на поверхні міцели і зберігали контакт із розчином. Виявлено, що молекули води проникають у змішані міцели на відстань до 0,3 – 1,0 нм від центру. Оцінено ступінь зв'язування протионів та встановлено його залежність від складу міцели. Виявлено, що він нелінійно зменшується зі зростанням вмісту ко-ПАР. Зокрема, за вмісту ко-ПАР до 34% зв'язування протионів було майже подібним до такого в чистих міцелах *n*-додецилсульфату натрію; значення поступово зменшувалося при підвищенні вмісту 1-бутанолу. Розглянуті декілька положень границі між зв'язаними та вільними протионами. Запропонований підхід може бути використаний для дослідження інших змішаних міцел.

Ключові слова: функція радіального розподілу, розмір міцели, гідратація, ступінь зв'язування протионів.

Kharkiv University Bulletin. Chemical Series. Issue 38 (61), 2022

ОСОБЛИВОСТІ СОЛЬВАТОХРОМІЇ 4-[[[(2,4-ДИНІТРОФЕНІЛ)МЕТИЛЕН]-ІМІНО-2,6-ДИФЕНІЛ]ФЕНОЛА ТА БАРВНИКА РАЙХАРДТА. DFT РОЗРАХУНКИ**Є. О. Сергєєва^a, А. Б. Захаров^b, С. М. Кійко^c**

Харківський національний університет імені В. Н. Каразіна, майдан Свободи 4, Харків, 61022, Україна

a) ✉ y.o.serhieieva@karazin.ua <https://orcid.org/0009-0002-9929-563X>b) ✉ abzakharov@karazin.ua <https://orcid.org/0000-0002-9120-8469>c) ✉ s.m.kiyko@karazin.ua <https://orcid.org/0000-0002-1161-5866>

Одним із актуальних напрямків розвитку сучасної фізичної хімії є розробка сенсорних пристроїв та молекулярних зондів для дослідження різних властивостей розчинів, колоїдних систем та біологічних об'єктів. До останніх відносяться сольватохромні барвники, які завдяки класичним роботам Райхардта, знайшли широке застосування для кількісної оцінки сольватуючої здатності індивідуальних та, в меншій мірі, змішаних розчинників різної природи.

Різна поведінка 4-[[[(2,4-динітрофеніл)метилен]іміно-2,6-дифеніл]фенола та барвника Райхардта в чистій воді та змішаних водно-органічних розчинниках, при зміні їх складу, вказує на те, що їх електронна будова зазнає принципової зміни при переході з основного у перший збуджений стан.

Метою роботи було дослідження та порівняння структури **HOMO** та **LUMO** барвника 4-[[[(2,4-динітрофеніл)метилен]іміно-2,6-дифеніл]фенола та стандартного бетаїнового барвника Райхардта з використанням стаціонарної та залежної від часу теорії функціонала густини (**DFT**).

Доведено, що барвник 4-[[[(2,4-динітрофеніл)метилен]іміно-2,6-дифеніл]фенол має два активних взаємозамінних сольватаційних центра й через це має відмінний механізм сольватації, принаймні у водному розчині, в порівнянні з барвником Райхардта, що повинно проявитися при його сольватації також в змішаних водно-органічних розчинниках з високим вмістом води в них.

Ключові слова: сольватохромія, сольватохромний барвник, полярність розчинника, квантово-хімічні розрахунки, теорія функціонала густини (**DFT**).

Вступ

Протягом останніх 40 років в дослідженні сольватохромних явищ перевага віддавалася класичному сольватохромному барвнику Райхардта (*N*-феноксіпіридинієвому бетаїну, 2,6-дифеніл-4-(2,4,6-трифенілпіридиній-1-іл) фенолят - далі **Ind 1**) або його похідним [1] та інших дослідників [2-9]. Використання інших сольватохромних барвників носило епізодичний характер. Це було зумовлено тим, що **Ind 1** проявляє сольватохромні ефекти в широкій області видимого та **УФ** спектра (від 453 до 925 нм) при переході із основного в збуджений стан, що пов'язано зі зменшенням дипольного моменту молекули індикатора на 22 D, що дозволяє навіть візуально спостерігати ці ефекти у видимій області спектру. [10,11] Така зміна електронної будови **Ind 1** в процесі спектрального переходу може бути представлена наступним чином (Рис.1) [1]

Даний барвник має досить зручну особливість сольватації для співставлення сольватуючої здатності (полярності) чистих розчинників в термінах емпіричних параметрів $E_T(30)$, а саме **Ind 1** сольватується переважно по фенолятному атому кисню, оскільки універсальні взаємодії вносять на порядок менший вклад в ентальпію сольватації, в порівнянні з іон-дипольною взаємодією, на що вказують лінійні кореляційні залежності $E_T(30)$ від акцепторного числа (A_N) та плавні параболічні залежності від ϵ для чистих протонних та апротонних розчинників (рис. 2) (значення $E_T(30)$, що використовуються в даній статті були отримані експериментально для підтвердження літературних значень [17] та для підтвердження методики виконання експерименту).

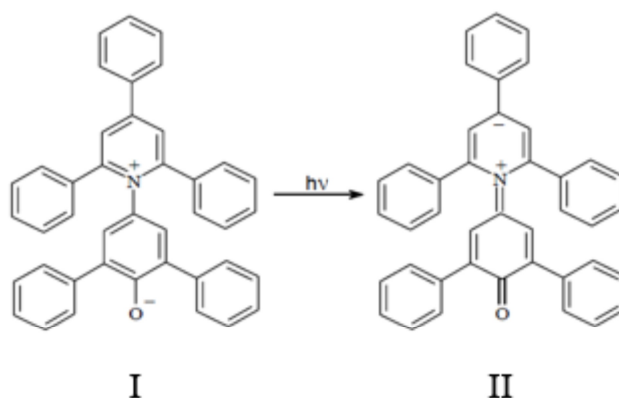
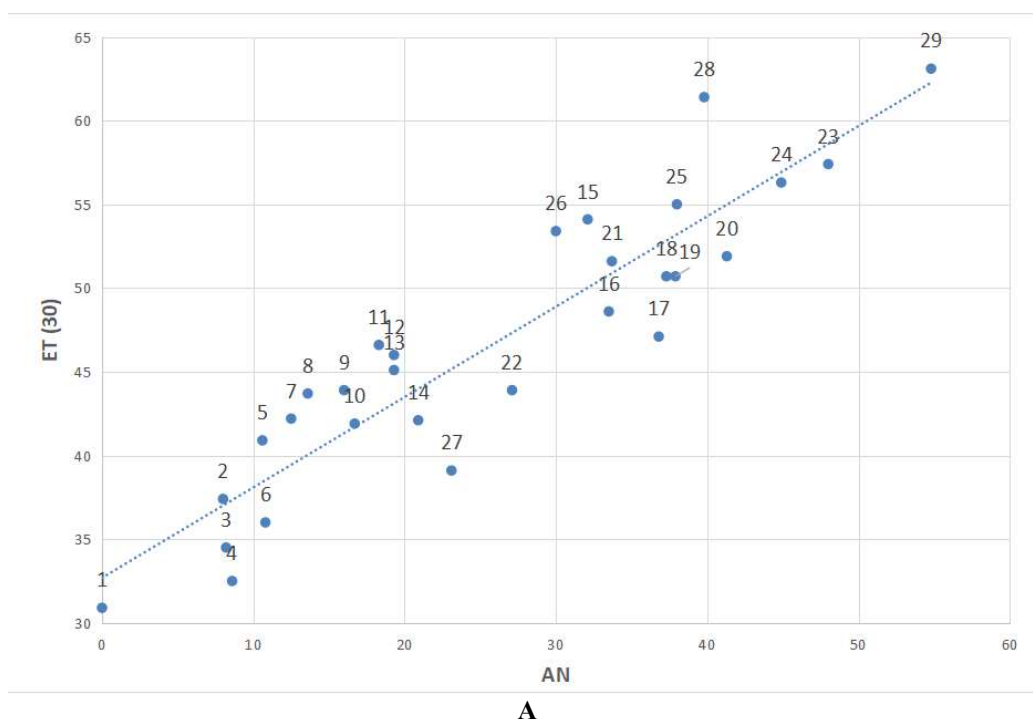
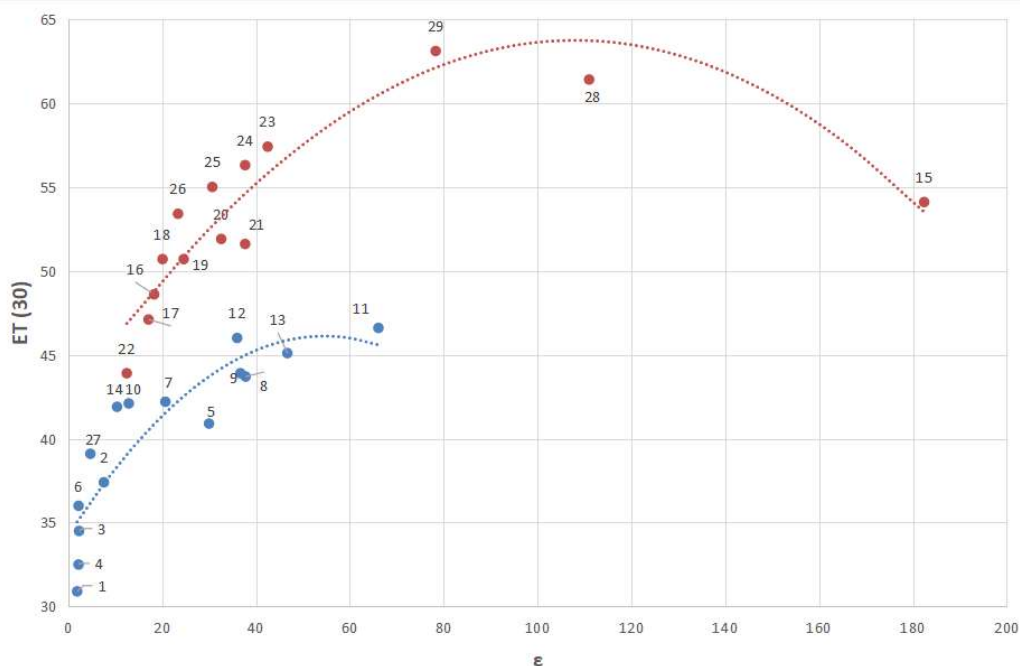


Рис. 1 Схематичне зображення будови електронного збудження в *Ind 1* [1]
 Fig. 1. Schematic representation of the structure of electronic excitation in *Ind 1* [1]

Залежність параметру $E_T(30)$ від акцепторного числа досліджена й доведена багатьма дослідницькими групами [12-15]. Також був показаний зв'язок параметру $E_T(30)$ з діелектричною проникністю. [16]

У зв'язку з тим, що спектральні дані *Ind 1* в змішаних водно-органічних розчинниках не несуть прямої інформації щодо специфіки міжмолекулярної взаємодії в змішаних водно-органічних розчинниках, були здійснені спроби синтезу нових сольватохромних індикаторів принаймні з двома різними сольватаційними центрами в основному та збудженому стані з приблизно тією ж областю прояву сольватохромії, як і у *Ind 1*.





В

Рис. 2 Залежності $E_T(30)$ від AN (А) [18,19] та ϵ (В) [20] чистих розчинників для *Ind 1*

Fig. 2 Dependences of $E_T(30)$ on AN (A) [18,19] and ϵ (B) [20] pure solvents for *Ind 1*

(1 – C_6H_4 , 2 – *THF*, 3 – C_6H_6 , 4 – CCl_4 , 5 – *HMPTA*, 6 – *DOx*, 7 – *AC*, 8 – *DMAA*, 9 – *DMFA*, 10 – *1,2-Dichloroethane*, 11 – *PC*, 12 – *ACN*, 13 – *DMSO*, 14 – *Ethylenediamine*, 15 – *n-MFA*, 16 – *i-PrOH*, 17 – *BuOH*, 18 – *n-PrOH*, 19 – *EtOH*, 20 – *MeOH*, 21 – *Ethanolamine*, 22 – *t-BuOH*, 23 – *Gly*, 24 – *EG*, 25 – *DEG*, 26 – *TEG*, 27 – $CHCl_3$, 28 – *FA*, 29 – H_2O)

Такий індикатор був синтезований у 2012 р. групою бразильських вчених [21] (4-[[[2,4-дінітрофеніл)метилєн]іміно-2,6-дифеніл]фєнол – далі *Ind 2*). Барвник може сольватуватися різними способами в залежності від складу змішаного водно-органічного розчинника, тобто переважно сольватується один з двох хромофорних центрів барвника. Він також проявляє сольватохромні ефекти в широкій області видимого та UV спектру (від 563 до 913 нм) (рис. 3).

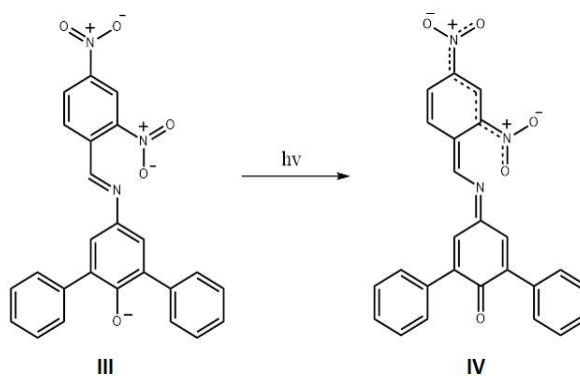


Рис. 3 Схематичне зображення будови електронного збудження в *Ind 2* [21]

Fig. 3. Schematic representation of the structure of electronic excitation in *Ind 2* [21]

У зв'язку з вищевказаним представляється цікавим проведення квантово-хімічних розрахунків бетаїнових барвників з метою детального дослідження їх електронних структур та подальшого обговорення.

Результати та їх обговорення

Завданням даної роботи було дослідження електронної будови обох барвників в основному стані. Оптимізація геометрії барвників та аналіз молекулярних орбіталей були проведені за допомогою пакета *Gaussian 09* [22] з використанням методу *DFT*, з функціоналом *m062x* та базисного набору *6-31+G(d, p)*. Результати наведені на рис. 4 – 5.

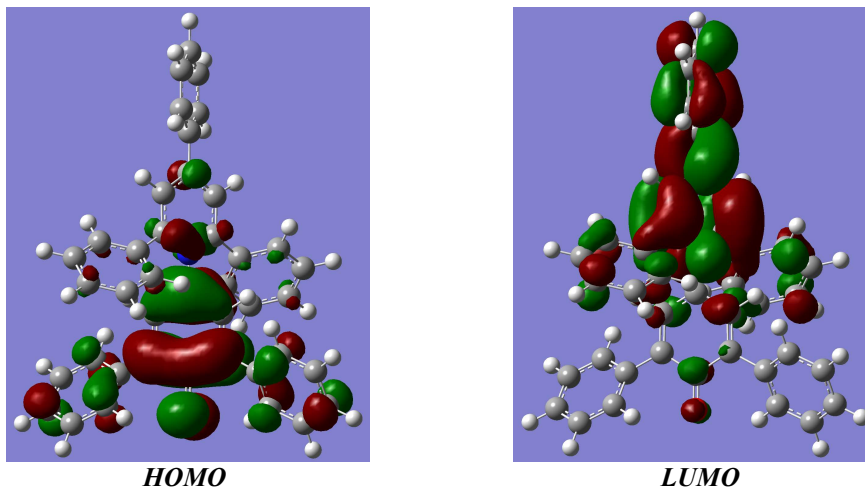


Рис. 4 Структури *HOMO* та *LUMO Ind 1*
Fig. 4 Structures *HOMO* and *LUMO* of *Ind 1*

Виходячи з результатів *DFT* розрахунків для основного стану *Ind 1* (структури *HOMO* та *LUMO* представлено на рис. 4) можна прийти до висновку, що орбіталі, які знаходяться поблизу щілини, локалізовані принципово на різних фрагментах молекули (*HOMO* переважно локалізована на «донорному» фенолятному фрагменті, в той час як *LUMO* в більшій мірі розташована на «акцепторній» частині барвника). Розмір щілини *HOMO-LUMO* складає 3.7 еВ. Згідно з нашими результатами *TD-DFT* розрахунків, перехід у збуджений стан реалізовано домінуючою конфігурацією *HOMO-LUMO*. Виходячи з цього, збудження відповідає внутрішньомолекулярному переносу заряду між донорним та акцепторним фрагментами, що тягне за собою значну зміну зарядового розподілу (а, як результат, до помітної різниці в дипольних моментах основного та збудженого станів), що й підтверджується експериментально у вигляді інтенсивної сольватохромії.

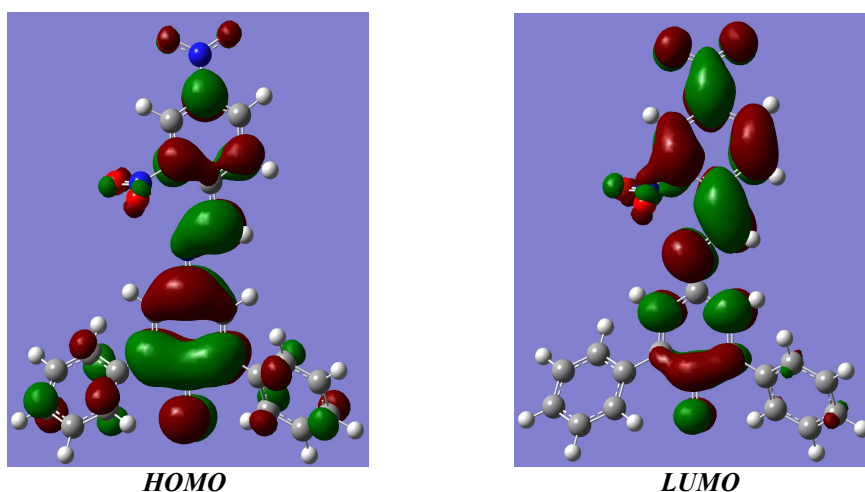
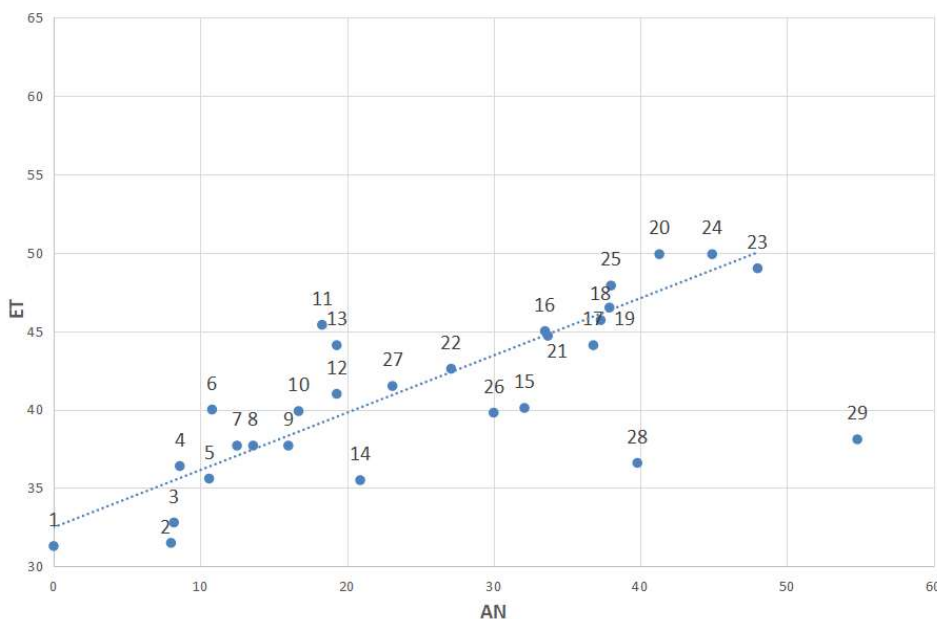


Рис. 5 Структури *HOMO* та *LUMO Ind 2*
Fig. 5 Structures *HOMO* and *LUMO* of *Ind 2*

На рис. 5, аналогічно випадку з *Ind 1*, представлено структури *HOMO* та *LUMO* для *Ind 2*. На відміну від *Ind 1* аналіз молекулярних орбіталей не виявив таких значних просторових розділень *HOMO* та *LUMO* між донорним та акцепторним фрагментами молекули (розмір щілини складає 3.25 eV). Слід відзначити, що верхня заповнена *MO* все ж таки має тенденцію до локалізації на фенолятному фрагменті, а нижня вакантна – на нітрогенвмістному фрагменті. Так само як і у випадку з *Ind 1*, перехід у збуджений стан реалізовано домінуючою конфігурацією *HOMO* → *LUMO*. У зв'язку з вищевказаним, не варто очікувати принципових відмінностей у зарядовому розподілі в основному та збудженому станах. Вірогідно, цей фактор вказує на той факт, що переважно фенолятний або нітрогенвмістний фрагменти в процесі сольватації можуть взаємодіяти з молекулами розчинника в залежності від його природи і лінійних розмірів його молекул. Слід зазначити, що розрахунки проводились без урахування специфічної взаємодії молекул індикатора та розчинника, що в свою чергу може значно вплинути на оптимальну структуру барвника за умов сольватації, і, як слід, на зарядовий розподіл.

Значна відмінність в сольватації *Ind 1* та *Ind 2* полягає в різній кількості сольватаційних центрів (один центр для *Ind 1* та два центри для *Ind 2*) – вірогідно, за умов відсутності стеричного затруднення в доступі до атома азоту в *Ind 1* (за рахунок фенільних кілець), для досліджуваних барвників могли б спостерігатись схожі особливості сольватації.

На основі цього можна зробити припущення, що стабілізація основного стану цих барвників буде змінюватися симбатно при переході від одного індивідуального розчинника до іншого. Це припущення повністю підтверджується залежностями параметрів $E_T(30)$ і E_T від AN та ϵ для більшості досліджених чистих розчинників різної природи, які наведені на рис. 2 і рис. 6, відповідно (значення E_T , що використовуються в даній статті були отримані експериментально для підтвердження літературних значень [20] та для підтвердження методики виконання експерименту).



А

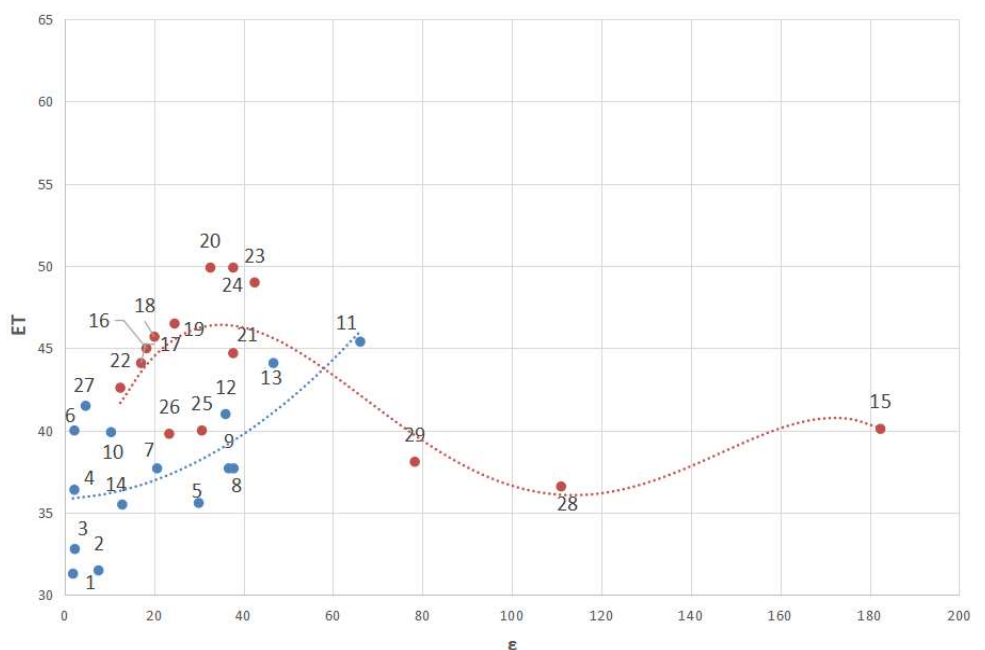

В

 Рис. 6 Залежності E_T від AN (А) [18,19] та ϵ (В) [20] чистих розчинників для **Ind 2**

 Fig. 6 Dependences of E_T on AN (A) [18,19] and ϵ (B) [20] pure solvents for **Ind 2**

(1 – C_6H_{14} , 2 – THF , 3 – C_6H_6 , 4 – CCl_4 , 5 – $HMPTA$, 6 – DOx , 7 – AC , 8 – $DMAA$, 9 – $DMFA$, 10 – *1,2-Dichloroethane*, 11 – PC , 12 – ACN , 13 – $DMSO$, 14 – *Ethylenediamine*, 15 – *n-MFA*, 16 – *i-PrOH*, 17 – *BuOH*, 18 – *n-PrOH*, 19 – $EtOH$, 20 – $MeOH$, 21 – *Ethanolamine*, 22 – *t-BuOH*, 23 – Gly , 24 – EG , 25 – DEG , 26 – TEG , 27 – $CHCl_3$, 28 – FA , 29 – H_2O)

Виняток складають тільки три розчинники: як ми можемо спостерігати, для **Ind 2** параметр E_T для H_2O сильно відрізняється від аналогічного параметру для **Ind 1**; також видно, що завдяки високим параметрам AN та ϵ із залежностей значно виділяються також FA та $n-MFA$, знову ж таки, для **Ind 2** у порівнянні з **Ind 1**. Така поведінка **Ind 2** в цих розчинниках може бути зумовлена, виключно, зміною механізму його сольватації, а саме сольватацією (стабілізацією) молекул **Ind 2** при певних умовах у збудженому стані. Що може вказувати на це?

По-перше, це структури **HOMO** та **LUMO** досліджуваних молекул. У випадку **Ind 1**, незважаючи на те, що в **HOMO** та **LUMO** індикатор акцентується на різних фрагментах, при використанні різних полярних молекул розчинників ми спостерігаємо, що нітрогенвмістний фрагмент, фактично, не приймає участі в сольватації. У аналогічній ситуації з **Ind 2**, при використанні різних полярних молекул розчинників ми можемо припустити, що барвник сольватується як по фенолятному фрагменту так і по нітрогенвмістному фрагменту (доступний для сольватації полярними молекулами, особливо, якщо ті здатні утворювати водневі зв'язки).

По-друге, як було сказано вище, із залежностей випадають тільки розчинники (рис. 6), які мають малі лінійні розміри, але при цьому характеризуються високими значеннями AN та ϵ . Оскільки, як видно з рис. 3, в збудженому стані **Ind 2** електронна густина (один e^-) приблизно рівномірно розподілена по ланцюжку із атомів **O-N-C-C-C-N-O**, то розмір сольватуючої молекули буде відігравати значну роль.

Це в першу чергу зумовлено малою об'ємною густиною електричного заряду на кожному атомі ланцюжку, що для стабілізації збудженого стану потребує утворення великої кількості слабких водневих зв'язків. При інших умовах більш енергетично вигіднішою буде сольватація **Ind 2** в основному стані з утворенням максимум 2-3 більш міцних водневих зв'язків за рахунок локалізації електрона на фенолятному атомі кисню. Останнє і спостерігається для **Ind 1**.

Висновки

Виходячи з результатів *DFT* розрахунків доведено, що *Ind 2* має два активних сольватаційні центри й через це має відмінний механізм сольватації від *Ind 1* в залежності від природи розчинника та лінійних розмірів його молекули.

Висловлене припущення, що стабілізація основного стану обох барвників має змінюватися симбатно при переході від одного індивідуального розчинника до іншого.

Встановлені деякі розчинники, які, вірогідно, можуть змушувати *Ind 2* працювати по іншому механізмі сольватації, відмінному від *Ind 1*.

Список літератури

1. Dimrot K.; Rneichard Ch.; Sripmann Th.; Bohlmann F. Uber Pyridinium-N-phenol-betaine und ihre Verwendung zur Charakterisierung der Polaritat von Losungsmitteln. Annalen der chemie band. 1963, 661, 1-37. <https://doi.org/10.1002/jlac.19636610102>.
2. Rettig W. Charge Separation in Excited States of Decoupled Systems TICT Compounds and Implications Regarding the Development New Laser Dyes and the Primary Processes of Vision and Photosynthesis. Angewandte Chemie International Edition in English. 1986, 25(11), 971-988. <https://doi.org/10.1002/anie.198609711>.
3. Samat A.; De Keukeleire D.; Guglielmetti R. Synthesis and Spectrokinetic Properties of Photochromic Spiropyrans. Bulletin des Societes Chimiques Belges. 1991, 100(9), 679-700. <https://doi.org/10.1002/bscb.19911000908>.
4. Shekhovtsov S.V.; Omelchenko I.V.; Dyakonenko V.V.; [et.al.] Synthesis and crystal structure determination of 2,6-di-tert-butyl-4-(2,4,6-triphenylpyridinium-1-yl)phenolate and its corresponding perchlorate salt. Dyes and Pigments. 2012, 92(3), 1394-1399. <http://doi.org/10.1016/j.dyepig.2011.06.029>.
5. Schramm A.D.S.; Nicoleti C.R.; Stock R.I.; Heying R.S.; Bortoluzzi A.J.; Machado V.G. Anionic optical devices based on 4-(nitrostyryl)phenols for the selective detection of fluoride in acetonitrile and cyanide in water. Sensors and Actuators B: Chemical. 2016, 240, 1036-1048. <https://doi.org/10.1016/j.snb.2016.09.052>.
6. Stock R.I.; Sandri Cr.; Rezende M.C.; Machado V.G. Solvatochromic behavior of substituted 4-(nitrostyryl)phenolate dyes in pure solvents and in binary solvent mixtures composed of water and alcohols. Journal of Molecular Liquids. 2018, 264, 327-336. <https://doi.org/10.1016/j.molliq.2018.05.042>.
7. Jeong J.; Su Min K.; Kumar R.S.; Mergu N.; Son Y.-A. Synthesis of novel betaine dyes for multi chromic sensors. Journal of Molecular Structure. 2019, 1187, 151-163. <https://doi.org/10.1016/j.-molstruc.2019.03.074>
8. Da Silva D.C.; Ricken I.; Silva M.A. do R.; Machado V.G. Solute-solvent and solvent-solvent interactions in the preferential solvation of Brooker's merocyanine in binary solvent mixtures. Journal of Physical Organic Chemistry. 2002; 15(7), 420-427. <https://doi.org/10.1002/poc.519>.
9. De Melo C.E.A.; Nicoleti C.R.; Nandi L.G.; [et.al.]. Solvatochromism of new substituted 4-[(E)-(4-nitrophenyl)diazenyl]phenolate dyes. Journal of Molecular Liquids. 2019, 301, 112330. <https://doi.org/10.1016/j.molliq.2019.112330>
10. Giusti L.A.; Marini V.G.; Machado V.G. Solvatochromic behavior of 1-(p-dimethylaminophenyl)-2-nitroethylen in 24 binary solvent mixtures composed of amides and hidoxilic solvents. Journal of molecular liquids. 2009, 150(1-3), 9-15. <https://doi.org/10.1016/j.molliq.2009.09.002>.
11. Machado V.G.; Stock R.I.; Reichardt C. Pyridinium N-Phenolate Betaine Dyes. Chemical Reviews. 2014, 114(20), 10429-10475. <https://doi.org/10.1021/cr5001157>.
12. Makitra R.G.; Pirig Y.N.; Kivelyuk R.B. Relation between parameter of electrophilicity of reichardt-dimroth et solvents and polarity and polarizability. Zhurnal Obshchei Khimii 1990, 60(10), 2209-2215.
13. Bekarek V.; Jurina J. A contribution to evaluation of influence of medium in electronic and infrared spectroscopy. Collect. Czech. Commun. 1982, 47, 1060-1068.
14. Fowler F.W.; Katritzky A.R.; Rutherford R.J. The correlation of solvent effects on physical and chemical properties. Journal of the Chemical Society B: Physical Organic. 1971, 460-469. <https://doi.org/10.1039/J29710000460>.

15. Zhang S.G.; Qi X.J.; Ma X.Y.; Lu L.J.; Deng Y.Q. Hydroxyl Ionic Liquids: The Differentiating Effect of Hydroxyl on Polarity due to Ionic Hydrogen Bonds between Hydroxyl and Anions. *The Journal of Physical Chemistry B*. 2010, 114(11), 3912-3920. <https://doi.org/10.1021/jp911430t>.
16. Schuster P.; Jakubetz W.; Marius W. Molecular models for the solvation of small ions and polar molecules. *Topics in Current Chemistry*. 1975, 60, 1-107. <https://doi.org/10.1007/BFb0045206>
17. Ceryn-Carrasco J.P.; Jacquemina D.; Laurence Chr.; Planchat A.; Reichardt Chr.; Sraidi Kh. Solvent polarity scales: determination of new $E_T(30)$ values for 84 organic solvents. *Journal of Physical Organic Chemistry*. 2014, 27(6), 512-518. <https://doi.org/10.1002/poc.3293>.
18. Evans L.T. A simple solubility theory combining solubility parameter and Lewis acid-base concepts. PhD. Thesis, Rochester Institute of Technology, 1988, 1-120.
19. Gutman V. Empirical parameters for donor and acceptor properties of solvent. *Electrochimica Acta*. 1976, 21, 661-670. [https://doi.org/10.1016/0013-4686\(76\)85034-7](https://doi.org/10.1016/0013-4686(76)85034-7)
20. Wohlfarth, Ch.: Landolt-Börnstein, New Series, Group IV, Physical Chemistry, Volume 17, Static Dielectric Constants of Pure Liquids and Binary Liquid Mixtures, Berlin: Springer-Verlag, 2008.
21. Nandi L.; Facin F.; Marini V.; [et.al.]. Nitro-Substituted 4-[(Phenylmethylene) imino]phenolates: Solvatochromism and Their Use as Solvatochromic Switches and as Probes for the Investigation of Preferential Solvation in Solvent Mixtures. *The Journal of Organic Chemistry*. 2012, 77(23), 10668-10679. <https://doi.org/10.1021/jo301890r>
22. Gaussian 09, Revision A.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2016.

Надійшло до редакції 14.06.2022

Прийнято до друку 12.09.2022

Y.O. Serhieieva, A.B. Zakharov, S.M. Kiyko. Peculiarities of solvatochromism of 4-[[2,4-dinitrophenyl)methylene]imino-2,6-diphenyl]phenol and Reichardt's dye. DFT calculations.

V. N. Karazin Kharkiv National University, faculty of chemistry, 4 Svobody sq., Kharkiv, 61022, Ukraine

One of the current directions of development of modern physical chemistry is the working out of sensor devices and molecular probes for the study of various properties of solutions, colloidal systems and biological objects. The latter include solvatochromic dyes, which, thanks to Reichardt's classic works, have found wide application for quantitative assessment of the solvating ability of individual and, to a lesser extent, mixed solvents of various nature.

The different behavior of Reichardt and 4-[[2,4-dinitrophenyl)methylene]imino-2,6-diphenyl]phenol dyes in pure water and mixed water-organic solvents, when their composition is changed, indicates that their electronic structure undergoes a fundamental change during the transition from the ground state to the first excited state.

The aim of the work was to study and compare the *HOMO* and *LUMO* structure of the standard Reichardt betaine dye and the 4-[[2,4-dinitrophenyl)methylene]imino-2,6-diphenyl]phenol dye using the stationary and time-dependent density functional theory (*DFT*).

It is proved that the 4-[[2,4-dinitrophenyl)methylene]imino-2,6-diphenyl]phenol dye has two active exchangeable solvation centers and therefore has an excellent solvation mechanism, at least in aqueous solution, compared to Reichardt dye, which should appear upon its solvation also in mixed water-organic solvents with a high water content in them.

Keywords: solvatochromia, solvatochromic dye, solvent polarity, quantum chemical calculation, density functional theory (*DFT*).

Kharkiv University Bulletin. Chemical Series. Issue 38 (61), 2022

ЕТИЧНІ НОРМИ ПУБЛІКАЦІЇ НАУКОВИХ РЕЗУЛЬТАТІВ ТА ЇХ ПОРУШЕННЯ.

Редакційна колегія робить все можливе для дотримання етичних норм, прийнятих міжнародним науковим товариством, і для запобігання будь-яких порушень цих норм. Така політика є важливою умовою плідної участі журналу в розвитку цілісної системи знань в галузі хімії та суміжних галузях. Діяльність редакційної колегії значною мірою спирається на рекомендації Комітету з етики наукових публікацій (Committee of Publication Ethics), а також на цінний досвід міжнародних журналів та видавництв. Подання статті на розгляд означає, що вона містить отримані авторами нові нетривіальні наукові результати, які раніше не були опубліковані. Кожну статтю рецензують щонайменше два експерти, які мають усі можливості вільно висловити мотивовані критичні зауваження щодо рівня та ясності представлення матеріалу, його відповідності профілю журналу, новизни та достовірності результатів. Рекомендації рецензентів є основою для прийняття остаточного рішення щодо публікації статті. Якщо статтю прийнято, вона розміщується у відкритому доступі; авторські права зберігаються за авторами. За наявності будь-яких конфліктів інтересів (фінансових, академічних, персональних та інших), учасники процесу рецензування мають сповістити редакційну колегію про це. Всі питання, пов'язані з можливим плагіатом або фальсифікацією результатів ретельно обговорюються редакційною колегією, рівно як спори щодо авторства та доцільність роздроблення результатів на невеличкі статті. Доведені плагіат чи фальсифікація результатів є підставами для безумовного відхилення статті.

STATEMENT ON THE PUBLICATION ETHICS AND MALPRACTICE. The Editorial Board has been doing its best to keep the ethical standards adopted by the world scientific community and to prevent the publication malpractice of any kind. This policy is considered to be an imperative condition for the fruitful contribution of the journal in the development of the modern network of knowledge in chemistry and boundary fields. The activity of the Editorial Board in this respect is based, in particular, on the recommendations of the Committee of Publication Ethics and valuable practice of world-leading journals and publishers. The submission of a manuscript implies that it contains new significant scientific results obtained by authors that were never published before. Each paper is peer reviewed by at least two independent experts who are completely free to express their motivated critical comments on the level of the research, its novelty, reliability, readability and relevance to the journal scope. These comments are the background for the final decision about the paper. Once the manuscript is accepted, it becomes the open-access paper, and the copyright remains with authors. All participants of the review process are strongly asked to disclose conflicts of interest of any kind (financial, academic, personal, etc.). Any indication of plagiarism or fraudulent research receives extremely serious attention from the side of the Editorial Board, as well as authorship disputes and groundless subdivision of the results into several small papers. Confirmed plagiarism or fraudulent research entail the categorical rejection of the manuscript.

ІНФОРМАЦІЯ ДЛЯ АВТОРІВ. Журнал публікує статті російською, англійською та українською мовами. До публікації приймаються: огляди (за погодженням з редколегією); оригінальні статті, обсяг 6-10 журнальних сторінок; короткі повідомлення, обсяг до 3 журнальних сторінок. Крім звичайного списку літератури, в статті обов'язково повинен бути другий список, всі посилання якого дані латиницею. Правила підготовки цього списку наведені в розділі «Транслітерація» на сайті журналу. Обидва списки повинні бути повністю ідентичні. При рецензуванні статей один з критеріїв - наявність посилань на публікації останніх років. Стаття обов'язково повинна містити резюме російською, українською та англійською мовами. У всіх трьох необхідно вказати назву статті, прізвища авторів і ключові слова. Орієнтовний обсяг резюме - 1800 знаків (без урахування заголовку і ключових слів). Редакція приймає електронний (MS Word) і два роздрукованих (для харків'ян) тексту рукопису. Адреси вказані в розділі «Контакти» на сайті журналу. Супровідний лист до статті, виправленої відповідно до зауважень рецензента, повинен містити відповіді на всі зауваження. Подається електронний і один роздрукований (для харків'ян) варіант. Рукописи, які пройшли рецензування, прийняті до публікації і оформлені відповідно до правил для авторів, приймаються у форматі doc (не docx) електронною поштою (chembull@karazin.ua). Роздрукований варіант не потрібен. Докладніша інформація розміщена на сайті журналу <http://chembull.univer.kharkov.ua>.

INFORMATION FOR AUTHORS. Papers in Ukrainian, Russian and English are published. These may be invited papers; review papers (require preliminary agreement with Editors); regular papers; brief communications. In preparing the manuscript it is mandatory to keep the statement on the publication ethics and malpractice, which can be found on the web-site and in each issue. The article should contain summaries in English, Russian, and Ukrainian. In all three it is necessary to indicate the title of the article, the names of the authors and the keywords. The approximate volume of summary is 1800 characters (excluding the title and key words). The help in translation is provided by request for foreign authors.. Any style of references is acceptable, but all references within the paper must be given in the same style. In addition, the second, transliterated, list of references is required if at least one original reference is given in Cyrillic. See section "Transliteration" of the web-site for details. Please use papers of previous issues as samples when prepare the manuscript. The MS Word format is used. Standard fonts (Times New Roman, Arial, Symbol) are preferable. Figures and diagrams are required in vector formats. Figure captions are given separately. All figures, tables and equations are numbered. Please use MS Equation Editor or MathType to prepare mathematical equations and ISIS Draw to prepare chemical formulas and equations. The decimal point (not coma) is accepted in the journal. Please avoid any kind of formatting when prepare the manuscript. Manuscripts may be submitted to the Editor-in-Chief via e-mail chembull@karazin.ua. For more detailed information see the journal web-site <http://chembull.univer.kharkov.ua>.

Наукове видання

Вісник
Харківського національного університету
імені В.Н. Каразіна

Серія «Хімія»
Вип. 38 (61)
Збірник наукових праць
Українською та англійською мовами.

Технічний редактор:
Д.О. Анохін

Підписано до друку «29» вересня 2022. Формат 60x84/8.
Ум.-друк. арк. 4,5 Обл.-вид. арк. 5,7.
Тираж 100 пр. Ціна договірна.

61022, Харків, майдан Свободи, 4
Харківський національний університет імені В.Н. Каразіна,
Видавництво Харківського національного університету імені В.Н. Каразіна

Надруковано: ХНУ імені В.Н. Каразіна
61022, м. Харків, майдан Свободи, 4.
Тел.: 705-24-32

Свідоцтво суб'єкта видавничої справи ДК № 3367 від 13.01.09