

06

Машинное обучение для определения распределения размеров металлических и фармакологических субмикронных частиц на основе индикатрис и экстинкций рассеянного света

© В.В. Ровенко, Т.А. Семенов, Е.О. Епифанов, Г.В. Мишаков, Н.В. Минаев,
А.А. Гарматина, Е.И. Мареев

Институт фотонных технологий Курчатовского комплекса кристаллографии и фотоники НИЦ „Курчатовский институт“, Москва, Россия

E-mail: rovenko.vladimir@physics.msu.ru

Поступило в Редакцию 12 января 2026 г.

В окончательной редакции 2 апреля 2026 г.

Принято к публикации 7 апреля 2026 г.

Предложена методика определения распределения субмикронных частиц по размерам (от 1 до 4000 nm), основанная на обучении сверточной нейронной сети по индикатрисам и экстинкциям рассеянного света. Обучение проводилось на синтезированных в рамках теории Ми матрицах оптического отклика для металлических (золото) и фармакологических (ибупрофен) частиц с учетом их комплексных показателей преломления. Предложенная нейросетевая модель позволяет восстанавливать параметры логнормального распределения частиц по размерам за время менее 1 s на графическом процессоре, обеспечивая среднее расхождение с данными сканирующей электронной микроскопии на уровне 30 %, что сопоставимо с точностью „ручного“ анализа при значительном сокращении временных затрат.

Ключевые слова: теория Ми, субмикронные частицы, светорассеяние, экстинкция, индикатриса рассеяния, машинное обучение.

DOI: 10.61011/PJTF.2026.13.63146.20621

Быстрое и точное определение распределения субмикронных частиц по размерам является критически важной задачей для нанотехнологий, материаловедения, фармакологии и экологического мониторинга [1]. В частности, контроль размеров металлических (например, золотых) наночастиц необходим для обеспечения их функциональности в катализе, сенсорике и биомедицине [2], а в случае фармакологических субмикронных частиц (таких как ибупрофен) размер напрямую влияет на биодоступность, растворимость и эффективность лекарственных средств [3]. Хотя электронная микроскопия остается одним из основных методов анализа, она сопряжена с трудоемкой пробоподготовкой и значительными временными затратами. Перспективной альтернативой являются методы, основанные на рассеянии света, которые позволяют проводить неразрушающий и оперативный *in situ* анализ. Однако извлечение информации о размере частиц из экспериментально измеренных индикатрис и экстинкций рассеянного света представляет собой сложную обратную задачу. Традиционный подход, основанный на „ручном“ подборе решения прямой задачи, чрезвычайно трудоемок, особенно для полидисперсных систем. Также следует отметить, что развитие современных методик сбора данных позволяет достичь скорости сбора спектральных данных 100–1000 Hz, что делает невозможным их обработку в „ручном“ режиме. В связи с этим активно развиваются разработки автономного определения размеров нано-

микрочастиц с помощью данных оптического рассеяния, включая нейросетевые подходы [4–7]. Так, например, в [5,6] продемонстрирована возможность определения размера монодисперсных наночастиц по спектрам экстинкции с использованием нейронных сетей, однако задача восстановления полного распределения частиц по размерам в полидисперсных системах остается менее изученной. А в [7] предложен метод оценки параметров распределения по данным динамического светорассеяния, но его применение ограничено субмикронным диапазоном и требует априорной информации о форме распределения. В работе [8] представлен нейросетевой подход для обратного проектирования произвольных рассеивателей на основе разложения по сферическим гармоникам, позволяющий предсказывать мультипольные коэффициенты и структурные параметры одиночных частиц с заданными электромагнитными свойствами. Данный подход ориентирован на дизайн метаатомов и не решает задачу определения распределения размеров в полидисперсных ансамблях. Настоящая работа развивает эти подходы, предлагая гибридную методику, сочетающую строгие расчеты на основе теории Ми с глубоким обучением для восстановления параметров логнормального распределения частиц по размерам на основе как спектров экстинкции, так и индикатрис рассеяния. Спектры экстинкции чувствительны к размеру частиц в широком диапазоне длин волн, а индикатрисы рассеяния предоставляют информацию об анизотропии рассеяния, что значительно расширяет

технические возможности осуществления контроля за получением нано- и микрочастиц.

Апробация методики проведена в прикладных задачах, связанных с генерацией золотых наночастиц методом лазерной абляции и формированием частиц ибупрофена методом быстрого расширения сверхкритических растворов (описание экспериментальных установок приведено в [9]). Для автоматизации процесса определения размеров частиц был разработан гибридный подход, сочетающий строгие расчеты на основе теории Ми с алгоритмами машинного обучения. На первом этапе с помощью разработанного программного обеспечения решается прямая задача светорассеяния и генерируются обширные базы данных, связывающие параметры распределения частиц с их оптическим откликом. Для золотых частиц использовались дисперсионные зависимости из работы [10], значения показателя преломления для ибупрофена измерены нами на рефрактометре DR-M2/1550 (Atago). На втором этапе обученная на этих данных нейронная сеть решает обратную задачу — восстанавливает распределение частиц по размерам на основе экспериментально измеренных индикатрис рассеяния или спектров экстинкции. Ключевое преимущество подхода заключается в способности нейросетевой модели аппроксимировать сложные нелинейные зависимости, что обеспечивает ее значительное превосходство перед классическими методами минимизации, основанными на линейной аппроксимации. Таким образом, предложенный подход расширяет возможности точного контроля размеров наночастиц в различных технологических процессах.

Программный пакет реализует расчеты базы данных на основе теории Ми для сферических частиц. Рассматривается классическая задача о рассеянии плоской монохроматической электромагнитной волны с длиной волны λ_0 на сферической частице радиуса a с комплексным показателем преломления $m = p + ik$, находящейся в непоглощающей среде с показателем преломления n_m [11,12]. В расчетах используется безразмерный параметр размера x :

$$x = k_m a = \frac{2\pi n_m}{\lambda_0} a, \quad (1)$$

где k_m — волновое число в окружающей среде. Для ансамбля частиц с заданным распределением по размерам $p(a)$ матрица оптического отклика рассчитывается как свертка решения задачи Ми для монодисперсных частиц. При этом предполагается, что система является разреженной, и межчастичным взаимодействием можно пренебречь, что соответствует условиям проведенных экспериментов (низкая концентрация частиц в жидкой или газовой фазе):

$$M_{base}(z) = \sum_a T_{scat}(x, z) \cdot p(x). \quad (2)$$

Для спектров экстинкции параметр z представляет собой длину волны излучения $z = \lambda$, а

$T_{scat}(x, z) = Q_{ext}(x, \lambda)$ — фактор эффективности экстинкции. Для индикатрис рассеяния z представляет собой угол рассеяния $z = \theta$, а $T_{scat}(x, z) = S(x, \theta)$ — амплитудные функции рассеяния. Для сопоставления с реальным экспериментом рассчитывалась относительная величина рассеяния вперед и назад $Ratio(\theta) = M_{for}(\theta)/M_{back}(\pi - \theta)$. В качестве $p(a)$ выбрано логнормальное распределение, хорошо зарекомендовавшее себя для описания подобных систем [9,13] и обоснованное физикой процессов лазерной абляции и быстрого расширения сверхкритических растворов:

$$p(a) = \frac{1}{a\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln a - \mu)^2}{2\sigma^2}\right), \quad a > 0,$$

$$\mu = \ln\left(\frac{M^2}{\sqrt{V+M^2}}\right), \quad \sigma = \sqrt{\ln\left(\frac{V}{M^2} + 1\right)}. \quad (3)$$

Параметры μ и σ связаны со средним арифметическим размером M и дисперсией V . Алгоритм формирования базы данных состоит из следующих этапов. Сначала производится параллельный расчет $T_{scat}(x, z)$ для всех заданных параметров и сохранение полученных результатов. Затем полученные результаты используются для вычисления матриц оптического отклика по формуле (2). Для аугментации данных к полученным результатам добавляется гауссов шум, чтобы повысить устойчивость модели к экспериментальным погрешностям. В результате формируется обширная база данных, устанавливающая соответствие между параметрами логнормального распределения (μ, σ) и синтезированными матрицами оптического отклика. Для каждого материала (золото, ибупрофен) генерировалась отдельная база данных с учетом его оптических констант.

Для решения обратной задачи определения параметров распределения частиц (μ, σ) по данным светорассеяния была разработана и обучена сверточная нейронная сеть, архитектура которой представлена на рис. 1. Модель, реализованная на основе библиотеки PyTorch, решает задачу многомерной регрессии [14]. Ее целью является прогнозирование параметров логнормального распределения (μ, σ) по входному спектру экстинкции или индикатрисе рассеяния. Архитектура сети включает один слой конволюции и четыре полносвязных слоя с размерами 128, 64, 32 и 16 нейронов. Выходной слой содержит 2 нейрона (μ и σ). Дальнейшее увеличение количества слоев не приводит к росту точности модели; уменьшение числа полносвязных слоев или отказ от использования конволюционного слоя приводит к значительному (более чем в 2 раза) падению качества восстановления. Обучение проводилось на графическом ускорителе Nvidia RTX 3070 с использованием оптимизатора Adam (learning rate = 0.001) и функции потерь MSE. Размер батча составлял 64. Существенного влияния размера батча и скорости обучения на конечный результат не обнаружено. Входные данные предварительно нормированы и представлены в виде одномерных

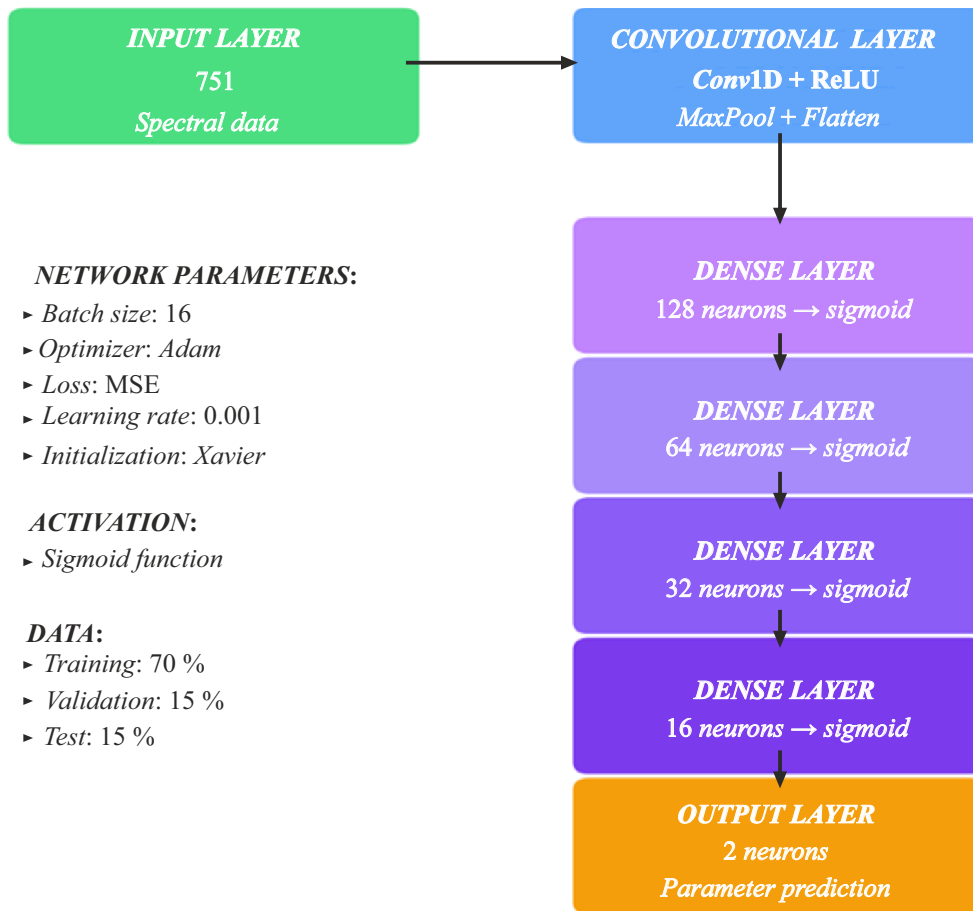


Рис. 1. Архитектура нейронной сети.

векторов фиксированной длины. Для спектров экстинкции: 751 точка, что соответствует диапазону длин волн 250–1000 nm с шагом 1 nm. Экспериментальные индикатрисы измерялись в восьми угловых положениях и интерполировались на 200 точек для подачи в нейросеть; в обучающей выборке использовались индикатрисы, рассчитанные с высоким угловым разрешением (1°).

Сходимость обучения контролировалась по динамике функции потерь на обучающей и валидационной выборках с функцией остановки обучения. Для валидации метода были сгенерированы базы данных, содержащие 100 000 уникальных распределений для каждого типа эксперимента. На рис. 2, *b* и 3, *a* представлено сравнение распределений, полученных с помощью сканирующей электронной микроскопии (СЭМ), с реконструированными классическим „ручным“ подбором и предложенной нейросетевой моделью для образцов золотых наночастиц (средний размер ~ 80 nm, получены лазерной абляцией в воде) и частиц ибупрофена (средний размер ~ 200 nm, получены методом RESS). На рис. 2, *a* и 3, *b* приведен результат моделирования спектра экстинкции для золота и индикатрисы рассеяния для ибупрофена на основании данных нейронной сети о распределении частиц, а также его сравнение с экспе-

риментом. Среднее расхождение между предсказанными и эталонными распределениями составило около 30 %, что в первую очередь объясняется отклонением формы реальных частиц от сферической, используемой в теории Ми. График функции потерь (рис. 4, *a*) показывает, что для достижения сходимости потребовалось около 165 эпох. Нейросетевая модель продемонстрировала низкую ошибку предсказания параметров ($MSE \sim 2.5 \cdot 10^{-3}$) на тестовой выборке (рис. 4, *a, b*), точность восстановления размеров слабо (в пределах 3 %, рис. 4, *b, c*) зависит от размеров частиц (в диапазоне 1 nm–1.5 μ m). Дополнительное тестирование на синтетических монодисперсных распределениях показало, что сеть корректно восстанавливает размер с погрешностью менее 5 % для частиц в указанном диапазоне чувствительности.

Таким образом, в работе предложен универсальный, точный и автоматизированный подход для определения распределения субмикронных частиц по размерам в реальном времени. Разработанный комплексный метод, сочетающий строгие расчеты на основе теории Ми с нейросетевым анализом данных, позволяет преодолеть ключевые ограничения традиционных методов, таких как трудоемкий „ручной“ подбор и линейная аппроксимация.

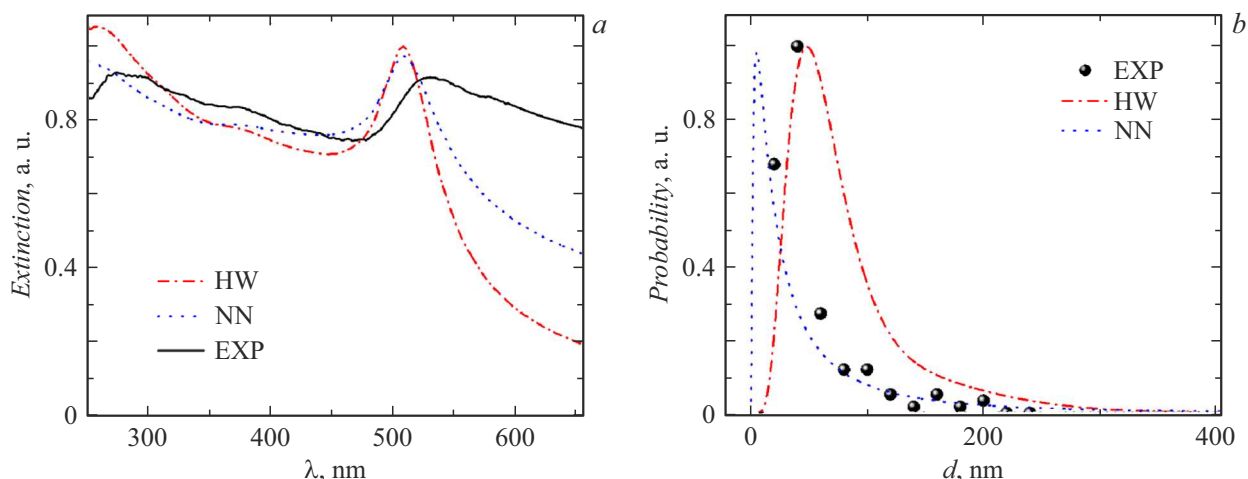


Рис. 2. Сравнение распределений частиц по размерам для образца золотых наночастиц, полученных методом сканирующей электронной микроскопии (EXP), „ручным“ подбором по оптическим данным (HW) и предсказанных нейросетью (NN). *a* — экспериментальные и модельные спектры экстинкции; *b* — вероятностные распределения частиц по размерам. Наблюдаемое расхождение составляет около 30% и обусловлено в первую очередь отклонением формы частиц от сферической.

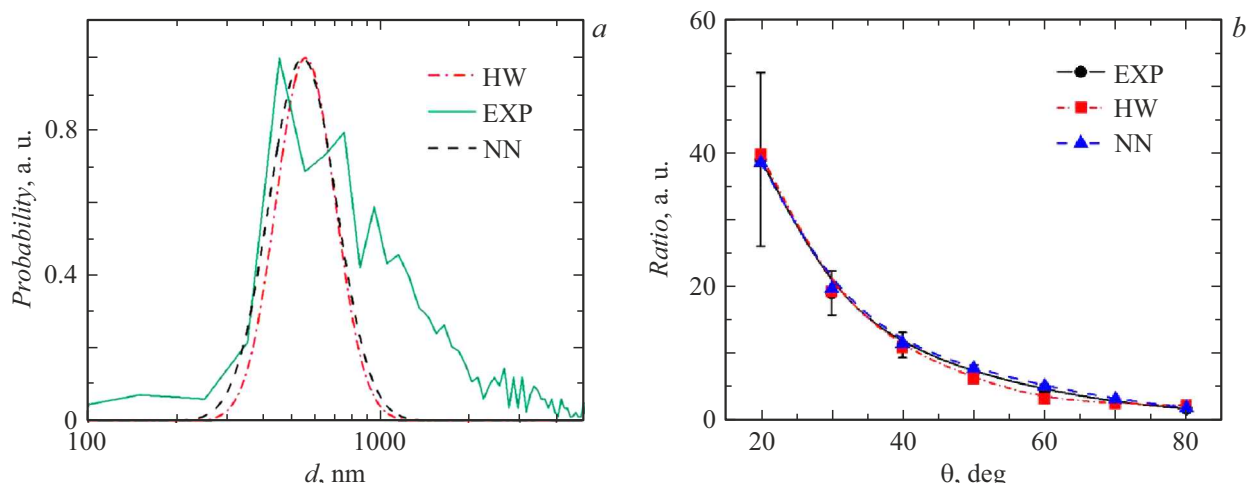


Рис. 3. Сравнение распределений частиц по размерам для образца частиц ибупрофена (метод RESS), полученных методом сканирующей электронной микроскопии (EXP), „ручным“ подбором (HW) и предсказанных нейросетью (NN). *a* — вероятностные распределения частиц по размерам; *b* — экспериментальные и модельные индикатрисы рассеяния. Показана сопоставимая с рис. 2 точность восстановления.

Главное преимущество подхода заключается в способности сверточной нейронной сети выявлять сложные нелинейные зависимости между оптическим откликом и параметрами распределения, что обеспечивает высокую точность прогнозирования на основе синтетических данных ($MSE \sim 2.5 \cdot 10^{-3}$) и быструю сходимость (около 165 эпох). Апробация методики на реальных экспериментальных данных подтвердила ее практическую состоятельность. Наблюдаемое расхождение с данными СЭМ на уровне 30% закономерно и в первую очередь связано с отклонением формы исследуемых частиц от сферической, что является фундаментальным ограничением используемой оптической модели, а не самого нейросетевого подхода. Полученные результаты демонстрируют потенциал метода для эффективного

контроля размеров частиц в различных технологических процессах.

Благодарности

Исследование размеров частиц с помощью СЭМ выполнено с использованием оборудования Центра коллективного пользования „Структурная диагностика материалов“ Курчатова комплекса кристаллографии и фотоники НИЦ „Курчатовский институт“.

Финансирование работы

Работа в части восстановления размеров наночастиц выполнена при поддержке гранта РФФ № 23-79-10188.

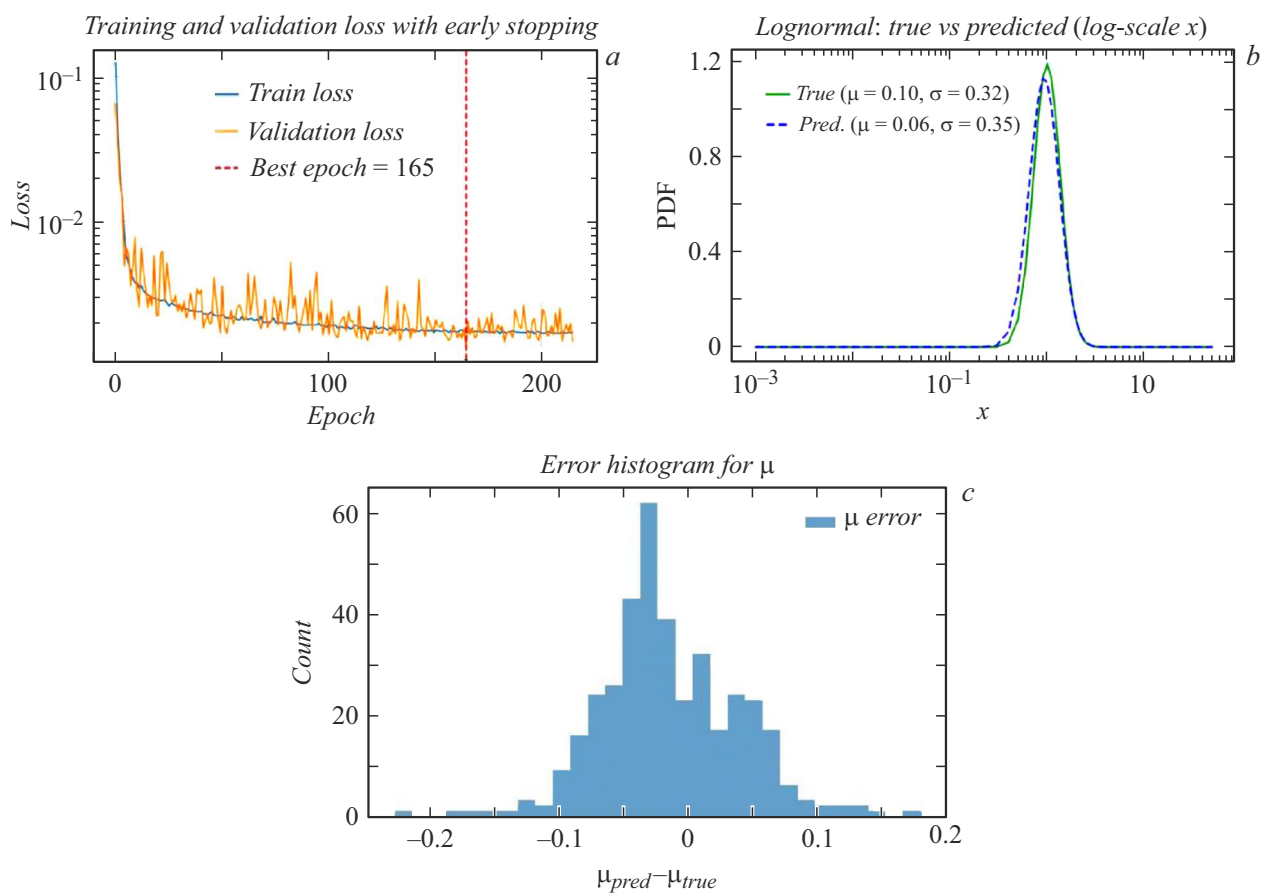


Рис. 4. Значения функции потерь на обучающей и валидационной выборках и сравнение предсказанного распределения с модельным. *a* — потери на обучающей и валидационной выборках в зависимости от количества эпох с указанием ранней остановки; *b* — вероятностные распределения частиц по размерам; *c* — гистограмма ошибок для параметра μ .

Конфликт интересов

Авторы заявляют, что у них нет конфликта интересов.

Список литературы

- [1] A.J. Hickey, C.-C. Huang, A. Misra, V. Balasubramanian, B.M. Moudgil, A.D. Rosato, L.M. Tavares, Q. Zhang, W.-N. Wang, *KONA Powder Part. J.*, **42**, 2026016 (2025). DOI: 10.14356/kona.2026016
- [2] S. Chakraborty, S. Mal, A. Halder, S. Das, K.K. Sen, A. Mukherjee, P. Roy, *Part. Sci. Technol.*, **42** (1), 145 (2024). DOI: 10.1080/02726351.2023.2211954
- [3] S. Jacob, F.S. Kather, S.H. Boddu, M. Attimarad, A.B. Nair, *Pharmaceutics*, **17** (1), 136 (2025). DOI: 10.3390/pharmaceutics17010136
- [4] T. Klinavicius, N. Khinevich, A. Tamuleviciene, L. Vidal, S. Tamulevicius, T. Tamulevicius, *J. Phys. Chem. C*, **128** (23), 9662 (2024). DOI: 10.1021/acs.jpcc.4c02459
- [5] A. Otsuki, H. Jang, *ChemEng.*, **6** (6), 92 (2022). DOI: 10.3390/chemengineering6060092
- [6] Y. Ren, J. Mao, H. Zhao, C. Zhou, X. Gong, Z. Rao, Q. Wang, Y. Zhang, *Adv. Meteorol.*, **2020**, 5074192 (2020). DOI: 10.1155/2020/5074192
- [7] M.A. Habitzreuter, M.C. Barbosa, *Phys. Rev. E*, **107** (3), 034601 (2023). DOI: 10.1103/PhysRevE.107.034601
- [8] J. Li, Z. Zhang, Y. Xu, S. Fu, J. Yang, Y. Wang, A.S. Shalin, Y. Qin, *Phys. Rev. Appl.*, **23** (1), 014001 (2025). DOI: 10.1103/PhysRevApplied.23.014001
- [9] E.O. Epifanov, N.M. Asharchuk, D.N. Hmelenin, I.V. Trofimov, A.O. Rybaltovskij, G.V. Mishakov, T.A. Semenov, N.V. Minaev, E.I. Mareev, *Russ. J. Phys. Chem. B*, **18** (8), 1948 (2024). DOI: 10.1134/S1990793124030047
- [10] P.B. Johnson, R.W. Christy, *Phys. Rev. B*, **6** (12), 4370 (1972). DOI: 10.1103/PhysRevB.6.4370
- [11] L.B. Scaffardi, J.O. Tocho, *Nanotechnology*, **17**, 1309 (2006). DOI: 10.1088/0957-4484/17/5/024
- [12] C.F. Bohren, D.R. Huffman, *Absorption and scattering of light by small particle* (Wiley-VCH, 1983), p. 102, 103.
- [13] H. Bagheri, H. Hashemipour, V. Ghalandari, S. Ghader, *Particuology*, **57**, 201 (2021). DOI: 10.1016/j.partic.2020.12.011
- [14] S.S.A. Zaidi, M.S. Ansari, A. Aslam, N. Kanwal, M. Asghar, B. Lee, *Comput. Biol. Med.*, **126**, 103514 (2022). DOI: 10.48550/arXiv.2104.11892