Радиология и радиобиология, радиационная безопасность

RADIOLOGY AND RADIOBIOLOGY, RADIATION SAFETY

УДК 543.429.3:539.16

ЭКСПЕРИМЕНТАЛЬНЫЙ АНАЛИЗ ЭФФЕКТИВНОСТИ МЕТОДА LS-SVM ДЛЯ СГЛАЖИВАНИЯ СПЕКТРОВ, ПОЛУЧЕННЫХ С ИСПОЛЬЗОВАНИЕМ МАЛОГАБАРИТНОГО СЦИНТИЛЛЯЦИОННОГО ДЕТЕКТОРА

А. Н. НИКИТИН¹⁾, Е. В. МИЩЕНКО¹⁾, Е. В. СОЛОНЕНКО¹⁾

¹⁾Институт радиобиологии Национальной академии наук Беларуси, ул. Федюнинского, 4, 246007, г. Гомель, Беларусь

Современные методы обработки спектрограмм гамма-излучения являются наиболее перспективными направлениями определения содержания радионуклидов в объектах окружающей среды. Особую проблему представляет задача выравнивания спектров при проведении измерений *in situ* из-за ограничений по времени набора и низкого отношения сигнал/шум. В настоящей работе приведена оценена эффективности использования метода регрессии на опорных векторах (LS-SVM) для сглаживания спектра, полученного с использованием сцинтилляционного NaI(Tl) детектора, в сравнении с методами скользящей средней и экспоненциального сглаживания. Оценка эффективности выравнивания спектра проводилась для всего энергетического диапазона измерения, а не только для области пика полного фотопоглощения квантов заданной энергии.

Образец цитирования:

Никитин АН, Мищенко ЕВ, Солоненко ЕВ. Экспериментальный анализ эффективности метода LS-SVM для сглаживания спектров, полученных с использованием малогабаритного сцинтилляционного детектора. Журнал Белорусского государственного университета. Экология. 2024;1:32–45. https://doi.org/10.46646/2521-683X/2024-1-32-45

Авторы:

Александр Николаевич Никитин – кандидат сельскохозяйственных наук, заместитель директора по научной работе. Егор Викторович Мищенко – научный сотрудник лаборатории радиоэкологии.

Евгения Викторовна Солоненко – научный сотрудник лаборатории радиоэкологии.

For citation:

Nikitin AN, Mischenko EV, Solonenko EV. Experimental analysis of the efficiency of gamma-ray spectrum smoothing with LS-SVM when using a compact scintillation detector. *Journal of the Belarusian State University. Ecology.* 2024;1:32–45. Russian.

https://doi.org//10.46646/2521-683X/2024-1-32-45

Authors:

Aleksander N. Nikitin, PhD (agriculture); deputy director for research.

nikitinale@gmail.com

Egor V. Mischenko, researcher at the laboratory of radioecology. *egormischenko@gmail.com Evgenia V. Solonenko*, researcher at the laboratory of radio-

ecology.

zhukvska ja.evgeni ja @gmail.com

Для анализа применялись полусинтетические спектры γ -излучения заданной продолжительности, сформированные поканальным суммированием радомизированно отобранных реальных спектров с переводом в скорость счета. Для оценки эффективности различных методов использованы полусинтетические спектры, сформированные для времени измерения 60, 300, 900, 1800, 3600, 7200 и 72000 с. Реальные спектры получены с использованием экспериментальной установки, состоящей из блока детектирования АТОМТЕХ, радионуклидного источника, компьютера со специализированным программным обеспечением и вспомогательных приспособлений. Выполненный анализ показал, что метод LS-SVM наиболее эффективен для сглаживания спектрограмм, полученных при времени измерения от $1 \cdot 10^3$ до $7 \cdot 10^3$ с. Отметим, что необходима раздельная оптимизация гиперпараметров модели сглаживания для трех частично перекрывающихся энергетических поддиапазонов. При более продолжительных измерениях применение испытанных методов сглаживания спектрограммы не приводит к снижению отношения сигнал/шум; при измерениях с коротким интервалом (до 100 с) наибольшей эффективностью обладает метод скользящей средней. Он позволяет повысить отношение сигнал / шум на 2–2,5 дБ. В диапазоне времени измерения 10^2-10^3 с наиболее сильное подавление статистического шума удается добиться с помощью экспоненциального сглаживания с $\alpha = 0,75$.

Ключевые слова: гамма-спектрометрия; обработка спектра; отношение сигнал-шум, регрессия на опорных векторах.

Благодарность. Исследования выполнены в рамках задания 3.05 НИР 4 ГПНИ «Природные ресурсы и окружающая среда».

EXPERIMENTAL ANALYSIS OF THE EFFICIENCY OF GAMMA-RAY SPECTRUM SMOOTHING WITH LS-SVM WHEN USING A COMPACT SCINTILLATION DETECTOR

A. N. NIKITIN^a, E. V. MISCHENKO^a, E. V. SOLONENKO^a

^aInstitute of Radiobiology of the National Academy of Sciences of Belarus, 4 Fiadzyuninskaga Street, Gomel 246007, Belarus Corresponding author: E. V. Mischenko (egormischenko@gmail.com)

Improving methods for processing gamma-ray spectrograms is a highly promising direction for further advancements in determining the content of radionuclides in environmental objects. A significant challenge lies in smoothing spectra during *in situ* measurements due to time constraints and low signal-to-noise ratio. This study evaluates the effectiveness of using Least Squares Support Vector Machine (LS-SVM) regression for smoothing spectra obtained with a NaI(TI) scintillation detector, comparing it with moving average and exponential smoothing methods. The evaluation of spectrum alignment effectiveness was conducted for the entire energy measurement range. Semi-synthetic gamma-ray spectra with measurement times 60, 300, 900, 1800, 3600, 7200, and 72000 seconds, generated by channel summing of randomly selected real spectra converted into count rates, were employed for analysis. Real spectra were acquired using an experimental setup consisting of an ATOMTEX detection unit, a radionuclide source, a computer with specialized software, and auxiliary devices. The analysis demonstrated that the LS-SVM method is most effective for smoothing spectrograms obtained at measurement times ranging from $1 \cdot 10^3$ to $7 \cdot 10^3$ seconds. Separate optimization of smoothing model hyperparameters is required for three partially overlapping energy subranges. For more longer measurements, applying tested spectrogram smoothing methods does not reduce the signal-to-noise ratio. For short intervals (up to 100 seconds), the moving average method exhibits the highest efficiency, allowing for a 2–2.5 dB improvement in the signal-to-noise ratio. In the measurement time range of 10^2 – 10^3 seconds, the most effective suppression of statistical noise is achieved with exponential smoothing with $\alpha = 0.75$.

Keywords: gamma-spectrometry; spectrum unfolding; signal-to-noise ratio, support vector regression

Acknowledgments. This study was supported by the Belarussian state scientific research program «Natural Resources and Environment» (2021–2025) task 3.05.4.

Введение

Спектрометрия гамма-излучения является мощным аналитическим методом, используемым в различных областях науки и техники для изучения свойств радиоактивных материалов и их взаимодействия с окружающей средой [1–4]. Картирование содержания радиоактивных изотопов в объектах окружающей среды с использованием полевой γ-спектрометрии широко используется не только для радиационного обследования территории и ликвидации радиационных аварий, но и для геологических исследований, поиска месторождений полезных ископаемых [5]. Развитию методов полевой спектрометрии способствует прогресс как в направлении приборостроения, так и в области методов обработки первичных данных. Согласно [6], основное влияние на точность результатов γ-спектрометрического анализа оказывают два источника ошибок: вариабельность параметров калибровки по энергии и статистический шум.

Первая проблема решается автоматической коррекцией усиления для каждого из кристаллов. Причина второй – вероятностная природа радиоактивного распада [7]. Проявлением статистического шума в спектрограмме γ-излучения является отклонение скорости счета в каналах реального спектра от ожидаемой скорости счета в каналах «идеального» спектра. По мнению экспертов [5], разработка методов повышения отношения сигнал/шум (SNR) в спектрограмме остается актуальной первоочередной задачей в области γ-спектрометрии. В настоящее время она решается за счет увеличения общего объема кристаллов детектора, что используется преимущественно при проведении измерений с использованием средств авиации, а также за счет новых методов выравнивания спектров при их обработке.

Актуальность поиска оптимального способа снижения статистического шума в спектре γ -излучения при проведении полевой спектрометрии обусловлена тем, что пространственное разрешение результата обратно пропорционально количеству точек измерения или скорости перемещения детектора. На практике количество точек измерения или скорость перемещения детектора ограничиваются временем, отведенным на проведение обследования заданной площади. Сохранение пространственного разрешения за счет уменьшения времени набора спектра сопровождается ростом статистического шума в спектрограмме [5]. Наличие метода сглаживания спектрограммы без искажения информативной составляющей (снижение вклада статистического шума) позволяет минимизировать время отбора отдельного спектра, повышая тем самым эффективность проведения полевой спектрометрии [5].

Спектр γ-излучения представляет собой сложную структуру с различной степенью ковариации между каналами [8; 9]. Отталкиваясь от этого положения, был предложен ряд методов снижения статистического шума (выравнивания спектра). Распространенными способами решения этой проблемы являются скользящая средняя и экспоненциальная скользящая средняя, сглаживание с использованием марковских цепей, Фурье-преобразование, вейвлет-преобразование и некоторые другие методы [10–13].

Увеличение вычислительной мощности мобильных компьютеров позволяет применять методы подавления статистического шума, основанные на обработке полного спектра. В настоящее время наиболее часто используются подходы, основанные на методе главных компонент. К их числу относятся MNF [14] и NASVD [15]. Недостатки таких методов – появление аномалий в выровненных спектрах и сильная зависимость результата от количества отбираемых компонент.

Альтернативным подходом к решению проблемы уменьшения статистического шума в спектрограмме с коротким временем измерения является ее представление с позиции особого случая аппроксимации функции с использованием методов машинного обучения [16]. Решение подобных задач состоит в обучении модели в направлении минимизации выбранной функции потерь. Для сглаживания спектрограммы модель каждый раз обучается по данным измерений и должна давать на выходе спектрограмму с уменьшенным вкладом статистического шума. В этой ситуации чрезвычайно важным этапом является правильный подбор гиперпараметров сглаживающей модели. Они должны обеспечивать сохранение значимой информации, но минимизировать амплитуду случайных отклонений, то есть построенная в последующем модель должна обладать оптимальной способностью к генерализации. Оптимизация гиперпараметров может быть осуществлена с использованием скорости счета в каналах эталонного спектра в качестве истинных значений и скорости счета в каналах на выходе сглаживающей модели в качестве предсказанных значений. В качестве функции потерь здесь обосновано использование среднеквадратичной ошибки.

Для шумоподавления в стохастических системах Sun J. и Zhou Y. [16] испытан метод регрессии на опорных векторах с наименьшими квадратами (LS-SVM). Метод опорных векторов позволяет проецировать мультиразмерное пространство параметров системы со сложной нелинейной зависимостью в другое мультиразмерное пространство, где зависимость между параметрами линейна, используя для этого набор нелинейных функций трансформации [17]. Зависимость между параметрами устанавливается в линейном пространстве. Испытания, выполненные Sun J. и Zhou Y. на синтетических временных рядах, свидетельствуют, что данный подход позволяет выделить значимую информацию и существенно снизить вклад статистического шума без предварительного знания о характеристиках системы.

Впервые использовать метод регрессии на несущих векторах для сглаживания спектрограмм у-излучения было предложено в 2017 г. Liu J, et al. [18]. Авторами были показаны перспективы данного метода обработки спектрограмм, получаемых при авиационной спектрометрии с использованием набора сцинтилляционных твердотельных детекторов большого объема.

Малогабаритные γ-спектрометрические комплексы на основе кристаллов NaI(Tl) получили широкое распространение в полевой спектрометрии. Они не требуют охлаждения, неприхотливы в обслуживании и хранении, обладают низким потреблением энергии. Современные интеллектуальные блоки детектирования на основе кристаллов NaI(Tl) включают в свой состав фотоумножитель, предусилитель, стабилизатор усиления и микроконтроллер для управления системой, формирования и хранения спектрограммы. Метод подавления шума на основе LS-SVM не был адаптирован (не подобраны оптимальные гиперпараметры модели) для подобных блоков детектирования. Насколько известно авторам, ранее не выполнялась оценка его эффективности для малогабаритных γ-спектрометрических комплексов на основе кристаллов NaI(Tl). Вместе с тем решение этой задачи важно для организации полевой γ-спектрометрии, исследования пространственного распределения радионуклидов в почве и других объектах окружающей среды.

Целью настоящей работы явилась адаптация метода регрессии на опорных векторах для сглаживания спектра γ-излучения, полученного с использованием малогабаритного интеллектуального блока детектирования на основе кристалла NaI(Tl) и сравнение его эффективности с наименее требовательными к вычислительным ресурсам методами – скользящей средней и экспоненциальной скользящей средней.

Следует отметить, что в полевой γ-спектрометрии могут представлять интерес не только области полного фотопоглощения квантов с заданной энергией, но и другие участки спектрограммы. В частности, соотношение между площадью пика и плато используется для оценки глубины залегания «центра масс» запаса радионуклида из выпадений [19]. Поэтому целесообразно провести оценку эффективности выравнивания для спектрограммы в целом, а не отдельных ее участков.

Материалы и методы исслдования

Экспериментальные работы проведены с использованием сцинтилляционного детектора на основе кристалла NaI(Tl) размером ø63×63 мм (196,2 см³) в составе интеллектуального блока детектирования Атомтех BDKG-07. Характеристики энергетической калибровки блока в период проведения измерений: граничная энергия – 3270 кэВ, 6,4 кэВ на один канал. Блок детектирования имеет систему встроенной светодиодной стабилизации измерительного тракта, предупреждающей дрейф по шкале энергии. При наборе спектрограмм блок детектирования без защиты располагался в закрытом помещении. На протяжении эксперимента температура в помещении составляла от 17,6 до 19,0 °C. В 0,1 м от него был размещен образцовый радионуклидный источник объемом 100 мл с удельным весом 1,02 г/см³ в геометрии ø70×26 мм, состав которого указан в табл. 1. Схема экспериментальной установки представлена на рис. 1.

Для управления блоком детектирования и сохранения результатов измерения использовано программное обеспечение ATAS версии 3.2 (Атомтех). Набор и сохранение спектров осуществлялись в автоматическом режиме циклических измерений. Время набора каждого спектра – 15 с, количество каналов – 512. Спектрограммы сохранялись в файлы формата SPE. Всего было набрано 8700 спектров. Перед обработкой спектры приведены к скорости счета в каналах (имп./сек.).

Обработку данных выполняли в среде *iPython* с использованием библиотек для обработки данных *Pandas, numpy, scipy*. Построение моделей для выравнивания спектрограмм осуществляли с использованием библиотеки *scikit-learn*. Для построения графиков использован инструментарий *matplotlib*.

Таблица 1

Состав образцового раствора радионуклидов по состоянию на момент проведения эксперимента

Table 1

| TI •.• •. | 1 0 1 0 | e 1º 1º 1 / /1 | 4 6 41 | • • |
|----------------------|------------------------|------------------------|---------------------|----------|
| The composition of t | the reference solution | of radionuclides at th | he moment of the ev | neriment |
| inc composition of t | inc reference solution | or rautomuchues at th | it moment of the ca | perment |

| Нуклид | Активность, Бк | Период полураспада |
|-------------------|----------------|--------------------|
| ⁵⁴ Mn | 64 | 312,34 сут. |
| ⁵⁷ Co | 25 | 271,79 сут. |
| ⁶⁰ Co | 152 | 5,27 лет |
| ⁶⁵ Zn | 7 | 244,26 сут. |
| ⁸⁸ Y | 2 | 106,65 сут. |
| ¹⁰⁹ Cd | 175 | 462,64 сут. |
| ¹¹³ Sn | 4 | 115 сут. |
| ¹³⁷ Cs | 286 | 30,07 лет |
| ¹³⁹ Ce | 2 | 137,64 сут. |
| ²⁴¹ Am | 863 | 432,27 лет |



Рис. 1. Схема экспериментальной установки: 1 – радионуклидный источник; 2 – штатив; 3 – опора; 4 – блок детектирования; 5 – блок сопряжения; 6 – компьютер со специализированным программным обеспечением

Fig. 1. Design of experimental set-up: 1 - radioisotope emitter; 2 - rack; 3 - support; 4 - detector unit; 5 - adapter; 6 - computer with specialized software

Для анализа использовались полусинтетические спектры γ-излучения заданной продолжительности (*t*), сформированные методом бутстрэппинга: рандомизированной выборкой соответствующего количества (*n*) 15-секундных спектров и расчета средней арифметической по каждому каналу. Данная процедура равносильна получению спектра, продолжительность набора которого равна 15 *n* с. Обоснованность данного подхода для формирования набора спектрограмм, используемых для машинного обучения, показана в [20]. Кроме исходных спектров с временем измерения 15 с, в анализе использованы полусинтетические спектры, сформированные для времени измерения 60, 300, 900, 1800, 3600, 7200 и 72000 с. Время измерения 15, 60, 300, 900, 1800 и 3600 с было выбрано на основании опыта авторов при проведении полевой спектрометрии, как наиболее часто используемые режимы сбора данных. Выбранные значения времени измерения отражают широкий диапазон условий получения спектрограмм, позволяя получить сравнительную характеристику эффективности испытываемых методов сглаживания спектров при различных отношениях сигнал/шум. Для анализа эффективности каждого из вариантов обработки было сформировано 1000 спектров.

Спектр, полученный посредством усреднения всей выборки (суммарное время набора 36,25 ч) с последующим пересчетом на скорость счета, был условно принят за эталонный, лишенный статистического шума (рис. 2). Он использован для оценки отношения сигнал / шум полусинтетических спектров до и после выравнивания. Для оценки адекватности использованного способа формирования эталонного спектра было произведено непрерывное измерение в аналогичных условиях на протяжении 10 ч (максимальная продолжительность, доступная в программном обеспечении ATAS) с последующим пересчетом на скорость счета. Среднеквадратичное отклонение между полусинтетическим эталонным и непрерывным спектрами составило 2,4·10⁻⁶, что свидетельствует о высокой степени их сходства и адекватности использованного метода для формирования полусинтетических спектров.



Рис. 2. Спектр ү-излучения, принятый за эталонный

Fig. 2. The reference γ -ray spectrum

Для оценки эффективности шумоподавления в спектре использовали меру ошибки L2, выражаемую в dB:

$$SNR = -10 \log_{10} \left(\frac{\|f_0 - \underline{f}\|^2}{\|f_0\|^2} \right), \tag{1}$$

где f_0 – эталонный спектр, с⁻¹; f – исходный спектр или спектр после выравнивания, с⁻¹. В данном исследовании рассматривается статистический шум во всем спектре γ -излучения, а под эффективностью шумоподавления подразумевается уменьшение статистического шума.

Сравнительный анализ статистического шума в спектрограмме ү-излучения имеет смысл для заданных условий измерений и диапазона регистрируемых энергий ү-квантов. Анализ полученных результатов для всего диапазона энергий детектора показывает, что отношение сигнал / шум в используемой экспериментальной установке имеет близкую к логарифмической зависимость от времени измерения (рис. 3). После 10⁴ с измерения рост SNR начинает замедляться сильнее. При увеличении времени измерения с 15 до 300 с SNR возрастает от 9,23 до 15,71 дБ. При 1-часовом измерении SNR достигает 21,05 дБ и приближается к 27 дБ при измерении, продолжающемся 24 ч.



Рис. 3. Зависимость отношения сигнал / шум от времени измерения. Область ошибок соответствует трехкратному среднеквадратичному отклонению SNR

Fig. 3. Dependence of the signal-to-noise ratio on the measurement time. The error region corresponds to 3-x standard deviation of the SNR

Результаты исследования и их обсуждение

В настоящей работе использован метод регрессии на несущих векторах (LS-SVM) с ядром преобразования к линейному пространству на основе радиальной базисная функция (RBF), испытание других вариантов ядер трансформации не производилось. Данная функция позволяет получить наилучшие результаты в различных приложениях метода несущих векторов [17]. Как было отмечено выше, эффективность выравнивания при использовании аппроксимирующих функций в значительной мере зависит от гиперпараметров модели машинного обучения, поэтому их подбор является важным этапом работы. Основными гиперпараметрами модели регрессии на опорных векторах являются: С – параметр регуляризации, чем он ниже, тем больше штрафуются сложность модели; γ – определяет строгость отбора опорных векторов; ϵ – расстояние от ограничивающей линии, проходящей через опорные вектора, до гиперплоскости.

Подбор гиперпараметров модели выравнивания осуществлялся методом Байесовой оптимизации. Критерием оптимизации служило максимальное значение отношения сигнал/шум. Оптимизация осуществлялась на спектрах с временем измерения 60 с. В данных условиях оптимальными параметрами оказались: $C = 251,4; \gamma = 0.0135, \epsilon = 0.0239.$

Проведенный анализ показал, что для спектров с продолжительностью набора 60 с, обработка с помощью LS-SVM повышает отношение сигнал/шум с 12,25 до 14,16 (табл. 2). Заметное улучшение данного показателя наблюдается также для наиболее коротких по продолжительности набора спектров (15 с). Для спектров с продолжительностью набора 300 с улучшение SNR незначительное. При увеличении продолжительности измерения применение метода LS-SVM вносит значительные искажение и ухудшает результат. Следует предположить, что оптимизация гиперпараметров модели LS-SVM на спектрограммах с более продолжительным временем набора позволит несколько исправить эту ситуацию. Но наибольший интерес представляет улучшение SNR для случаев с короткой продолжительностью набора.

Таблица 2

Отношение сигнал / шум при выравнивании спектра γ-излучения методами LS-SVM; средняя арифметическая ± 3 среднеквадратических отклонения

Table 2

| Время измерения, с | Без выравнивания | LS-SVM | SVRw, 60 c | SVRw, 900 c |
|--------------------|----------------------|------------------|----------------------|----------------------|
| 15 | $9,24 \pm 0,36$ | 11,78 ± 0,61 | $9,76 \pm 0,40$ | $9,55 \pm 0,42$ |
| 60 | $12,25 \pm 0,34$ | $14,16 \pm 0,54$ | $12,\!66 \pm 0,\!41$ | $12,55 \pm 0,45$ |
| 300 | $15,74 \pm 0,35$ | 16,13 ± 0,28 | $16,\!07\pm0,\!42$ | $16,\!07 \pm 0,\!40$ |
| 900 | $18,\!10 \pm 0,\!36$ | $16,65 \pm 0,15$ | $18,\!43\pm0,\!42$ | $18,\!42 \pm 0,\!45$ |
| 1800 | $19,60 \pm 0,35$ | 16,69 ± 0,12 | $19,83 \pm 0,37$ | $19,93 \pm 0,47$ |
| 3600 | 21,06 ± 0,36 | $16,73 \pm 0,07$ | 21,16 ± 0,41 | $21,\!28\pm0,\!40$ |
| 7200 | 22,48 ± 0,37 | 16,69 ± 0,06 | $22,\!30\pm0,\!31$ | $22,\!82\pm0,\!41$ |
| 72000 | $26,47 \pm 0,33$ | $16,64 \pm 0,04$ | $24,46 \pm 0,13$ | $25,33 \pm 0,17$ |

Signal-to-noise ratio after γ -ray spectrums smoothing by LS-SVM method; arithmetic mean ± 3 standard deviation

Примечание. LS-SVM – регрессия на опорных векторах с одной моделью для всего энергетического диапазона и оптимизацией гиперпараметров для времени измерения 60 с; SVRw, 60 с – LS-SVM с раздельными моделями для трех энергетических окон и оптимизацией гиперпараметров для времени измерения 60 с; SVRw, 900 с – LS-SVM с раздельными моделями для трех энергетических окон и оптимизацией гиперпараметров для времени измерения 900 с.

Notes. LS-SVM – support vector regression with one model for the entire energy range and optimization of hyperparameters for a measurement time of 60 s; SVRw, 60 c – LS-SVM with separate models for three energy windows and hyperparameter optimization for a measurement time of 60 s; SVRw, 900 s – LS-SVM R with separate models for three energy windows and hyperparameter optimization for measurement time of 900 s.

Результаты γ-спектрометрических измерений без использования защиты характеризуются сильно различающимися скоростями счета и дисперсией в далеко отстоящих друг от друга энергетических каналах. Liu Jun, et al. [21] пришли к выводу, что использование единой модели LS-SVM для всего энергетического диапазона в подобной ситуации не позволяет достичь желаемого результата. Подобный подход приводит к необходимости использования слишком большого количества опорных векторов и переобучению моделей. На основании анализа результатов γ-аэросъемки, авторами выделено три энергетических окна, в которых аппроксимация выравнивающими моделями производится раздельно. Чтобы предупредить скачки значений в области перекрытия окон, вызывающий неоправданный рост первой производной, использована специальная процедура сглаживания на заданной ширине перекрытия окон [21].

Нами выполнено испытание эффективности предложенной Liu Jun, et al. [21] трехоконной оптимизации для спектрограмм, полученных с помощью интеллектуального блока детектирования Атомтех BDKG-07. Подбор гиперпараметров моделей осуществлялся методом Байесовой оптимизации для времени измерения – 60 с (SVRw, 60 с) и 900 с (SVRw, 900 с). В рамках решения данной задачи также подобраны оптимальные диапазоны для каждого из трех окон и ширина их перекрытия. Оптимальные гиперпараметры моделей представлены в табл. 3.

Выполненная оценка показала, что трехоконная модель LS-SVM оптимизированная на спектрограммах с временем набора 60 с имеет преимущество перед единой моделью только на спектрограммах с временем измерения более 900 с (см. табл. 2). При этом, начиная с времени измерения 1800 с и более, ее использование приводит лишь к минимальному улучшению SNR, а с 7200 с и более – ухудшает данный показатель.

Примеры выравнивания спектрограмм с использованием трехоконной модели LS-SVM, оптимизированной на время измерения 60 с, приведены на рис. 4. Модель позволяет в значительной мере сгладить спектрограмму, приблизив скорость счета в каналах к эталонным значениям. При этом на ней могут возникать небольшие артефакты, наиболее заметные при коротком времени измерения. Выровненная спектрограмма с временем измерения 900 с мало отличаются от эталонной, при этом на реальной спектрограмме статистический шум хорошо заметен.

Таблица 3

Гиперпараметры моделей регрессии на опорных векторах для выравнивания спектров γ-излучения, полученных на детекторе NaI(TI). Индекс при параметре обозначает принадлежность к модели соответствующего окна

Table 3

| Параметр | SVRw, 60 c | SVRw, 900 c |
|---------------------------------|------------------------|------------------------|
| C ₁ | 16,7 | 300,0 |
| C ₂ | 302,9 | 15,4 |
| C ₃ | 234,8 | 186,8 |
| €1 | 3,17.10-5 | 1,05.10-4 |
| ε2 | 6,131·10 ⁻⁴ | 8,988·10 ⁻³ |
| €3 | 5,27.10-3 | 3,71.10-4 |
| γ1 | 0,429 | 0,773 |
| γ ₂ | 3,65.10-3 | 1,80.10-2 |
| γ ₃ | 1,03.10-3 | 3,62.10-2 |
| Ширина перекрывания окон | 6 | 8 |
| Граница между окнами 1 и 2 (w1) | 22 | 27 |
| Граница между окнами 2 и 3 (w2) | 299 | 270 |
| Ядро | RBF | RBF |

Hyperparameters of the Support Vector Regression Models for smoothing γ-ray spectrum from the NaI(Tl) detector. The index at the parameter indicates the belonging to the model of the corresponding window

Оптимизация гиперпараметров трехоконной модели LS-SVM на спектрограммах с временем измерения 900 с демонстрирует некоторое улучшение отношения сигнал/шум при времени измерения 1800–7200 с. При менее продолжительном измерении она уступает единой модели LS-SVM и трехоконной модели LS-SVM с оптимизацией на спектрограммах с временем измерения 60 с.

При времени измерения 72000 с ни один из испытанных методов не позволят улучшить показатель SNR, демонстрируемый на реальной спектрограмме без обработки.

Некоторые примеры выравнивания спектрограммы с использованием трехоконной модели LS-SVM с оптимизацией на спектрограммах с временем измерения 900 с представлены на рис. 5. Появление артефактов на спектрограммах с временем измерения 15 и 60 с для этой модели более характерно чем для предыдущей. Но при времени измерения 900 с скорость счета в каналах выровненной спектрограммы мало отличается от эталонных значений.

Использование моделей LS-SVM для выравнивания спектрограмм требует предварительной оптимизации гиперпараметров и обработки с использованием производительного процессора. Вместе с тем в практике у-спектрометрии широкое применение имеет использование более простых методов выравнивания спектрограмм. Наиболее простым методом является скользящая средняя. Данная функция реализована в большинстве программных пакетов для обработки спектров. Единственным параметром данного фильтра является ширина фильтра (окна усреднения). Результаты анализа показали (табл. 4), что, с точки зрения увеличения отношения сигнал / шум, данный прием оказывается эффективным лишь при коротком времени измерения (15–300 с). При времени измерения 900 с улучшение SNR по сравнению с необработанной спектрограммой несущественное, а при более продолжительных измерениях SNR после обработки методом скользящей средней ухудшается. Применение скользящей средней к спектрам, набранным за 15 с, позволяет достичь уровня SNR, характерного для 60-секундных измерений. Существенным недостатком скользящей средней является уширение пиков на спектре, что означает снижение разрешения. Данное явление отчетливо проявляется при продолжительности времени измерения более 5–15 мин. В этой ситуации применение скользящей средней снижает SNR, что делает использование данного фильтра не только бесполезным, но и ухудшающим качество результата.





Fig. 4. Examples of spectrograms smoothing using the three-window LS-SVM model with hyperparameters optimized on spectrograms with a measurement time of 60 s: 0 – spectrogram after alignment; 1 – reference spectrogram; 2 – original real spectrogram



Рис. 5. Пример выравнивания спектра с использованием трехоконной модели LS-SVM с оптимизацией гиперпараметров на спектрограммах с временем измерения 900 с: 0 – спектрограмма после выравнивания; 1 – эталонная спектрограмма; 2 – исходная реальная спектрограмма

Fig. 5. Examples of spectrograms smoothing using the three-window LS-SVM model with hyperparameters optimized on spectrograms with a measurement time of 900 s: 0 – spectrogram after alignment; 1 – reference spectrogram; 2 – original real spectrogram

Таблица 4

Отношение сигнал / шум при выравнивании спектра γ-излучения методом скользящей средней; средняя арифметическая ± 3 среднеквадратических отклонения, l – ширина фильтра, канал

Table 4

| Время измерения, с | Без выравнивания | 1 = 3 | 1 = 5 | 1 = 7 | 1 = 9 |
|--------------------|------------------|----------------------|----------------------|----------------------|----------------------|
| 15 | 9,24 ± 0,36 | $11{,}58\pm0{,}48$ | $12,\!39\pm0,\!58$ | 12,51 ± 0,61 | $12,20 \pm 0,62$ |
| 60 | $12,25 \pm 0,34$ | $14,\!44\pm0,\!48$ | 14,61 ± 0,46 | $13,\!93 \pm 0,\!46$ | $12,93 \pm 0,41$ |
| 300 | 15,74 ± 0,35 | 17,24 ± 0,39 | $15,\!94 \pm 0,\!30$ | $14,\!48 \pm 0,\!25$ | $13,\!17 \pm 0,\!20$ |
| 900 | 18,10 ± 0,36 | $18,53\pm0,30$ | $16,29 \pm 0,19$ | $14,61 \pm 0,16$ | 13,21 ± 0,11 |
| 1800 | 19,60 ± 0,35 | $18,\!99\pm0,\!24$ | $16,38 \pm 0,14$ | $14,62 \pm 0,11$ | $13,23 \pm 0,08$ |
| 3600 | 21,06 ± 0,36 | $19,\!28 \pm 0,\!17$ | $16,\!44 \pm 0,\!10$ | $14{,}64\pm0{,}08$ | $13,\!23 \pm 0,\!06$ |
| 7200 | $22,48 \pm 0,37$ | 19,45 ± 0,13 | $16,\!46 \pm 0,\!07$ | $14,\!65 \pm 0,\!05$ | $13,\!23 \pm 0,\!04$ |
| 72000 | $26,47 \pm 0,33$ | 19,61 ± 0,04 | $16,\!48 \pm 0,\!02$ | $14,65 \pm 0,02$ | $13,24 \pm 0,01$ |

Signal-to-noise ratio after γ-ray spectrum smoothing by the moving average method; arithmetic mean ± 3 standard deviation; l is the filter width, channel

Наиболее эффективная ширина фильтра скользящей средней снижается с 7 до 3 при увеличении времени измерении с 15 до 300 с. Обращает на себя внимание тот факт, что при использовании скользящей средней SNR при росте времени измерения увеличивается до определенного предела, после чего остается практически неизменным. Максимальное значение SNR уменьшается при увеличении ширины фильтра скользящей средней.

По сравнению со скользящей средней, единая для всего спектра модель LS-SVM не дает никаких преимуществ и ее использование не может быть оправдано. Трехоконная модель LS-SVM, оптимизированная на спектрограммах с временем измерения 900 с, демонстрирует более высокое качество выравнивания по сравнению со скользящей средней на спектрограммах с временем измерения 1800–7200 с.

Экспоненциальное сглаживание является нелинейным фильтром, представляющим собой скользящую среднюю с весами, экспоненциально убывающими при удалении от центра фильтра. Сила сглаживания определяется параметром α . Теоретическим преимуществом экспоненциального сглаживания перед скользящей средней является то, что оно должно меньше искажать форму пиков и не так сильно ухудшать разрешение.

В табл. 5 представлена оценка качества сглаживания спектра γ-излучения с α равной 0,25, 0,5 и 0,75 при ширине фильтра равной 5 каналам.

Существенное увеличение показателя SNR относительно необработанной спектрограммы достигается при использовании экспоненциального фильтра с $\alpha = 0,25$ при времени измерения 15 и 60 с, а при использовании экспоненциального фильтра с $\alpha = 0,75$ – при времени измерения 300 с. Экспоненциальный фильтр имеет незначительные преимущества по сравнению со скользящей средней при времени измерения 60 ($\alpha = 0,25$), 900 ($\alpha = 0,75$) и 1800 ($\alpha = 0,75$) с. При времени измерения 15 с сохраняется преимущество скользящей средней с шириной окна равной 7 каналам.

Как и в случае со скользящей средней, трехоконная модель LS-SVM превосходит экспоненциальный фильтр при времени измерении 1800, 3600 и 7200 с. При продолжительности измерений 72000 с ни один из испытанных методов сглаживания спектра не улучшает SNR.

Дополнительно проведен более подробный анализ зависимости эффективности выравнивания спектров γ -излучения с помощью испытываемых методов от продолжительности измерения (рис. 6). Полученные результаты показывают, что испытанные методы подавления статистического шума в спектре γ -излучения позволяют улучшить SNR при времени измерения до 10⁴ с. При более продолжительных измерениях испытанные методы ухудшают SNR. Наиболее сильное улучшение отношения сигнал / шум достигается при непродолжительном времени измерения 1–5·10² с.

Таблица 5

Отношение сигнал / шум при выравнивании спектра γ-излучения с помощью экспоненциального фильтра шириной 5 каналов с различным параметром α

Table 5

| Время измерения, с | Без выравнивания | α = 0,25 | $\alpha = 0,5$ | $\alpha = 0,75$ |
|--------------------|------------------|------------------|----------------------|----------------------|
| 15 | 9,24 ± 0,36 | 12,41 ± 0,59 | $12,08 \pm 0,53$ | $11,05 \pm 0,42$ |
| 60 | $12,25 \pm 0,34$ | $14,82 \pm 0,51$ | $14,74 \pm 0,50$ | $13,97 \pm 0,43$ |
| 300 | $15,74 \pm 0,35$ | $16,47 \pm 0,34$ | 17,05 ± 0,38 | 17,16 ± 0,39 |
| 900 | $18,10 \pm 0,36$ | $16,92 \pm 0,22$ | $17,\!84 \pm 0,\!29$ | $18,90 \pm 0,32$ |
| 1800 | 19,60 ± 0,35 | 17,06 ± 0,16 | $18,11 \pm 0,21$ | 19,67 ± 0,30 |
| 3600 | 21,06 ± 0,36 | 17,12 ± 0,12 | $18,25 \pm 0,16$ | $20,\!19 \pm 0,\!24$ |
| 7200 | $22,48 \pm 0,37$ | 17,16 ± 0,08 | $18,34 \pm 0,11$ | 20,51 ± 0,19 |
| 72000 | $26,47 \pm 0,33$ | $17,20 \pm 0,03$ | $18,\!41 \pm 0,\!04$ | $20,\!86 \pm 0,\!07$ |

Signal-to-noise ratio after γ-ray spectrum smoothing using an exponential filter with width of 5 channels and different parameter α



Примечание. Условные обозначения: SVR – регрессия на опорных векторах – с одной моделью для всего энергетического диапазона и оптимизацией гиперпараметров для времени измерения 60 с; SVRw 900 с – LS-SVM с раздельными моделями для трех энергетических окон и оптимизацией гиперпараметров для времени измерения 900 с.

Note. Conventional designations: SVR – Support Vector Regression – with a single model for the entire energy range and hyperparameter optimization for a measurement time of 60 seconds; SVRw 900 c – LS-SVM with separate models for three energy windows and hyperparameter optimization for a measurement time of 900 seconds.

Рис. 6. Эффективность скользящей средней и регрессии на опорных векторах для выравнивания спектров у-излучения при различной продолжительности измерения

Fig. 6. Efficiency of the moving average and support vector regression methods for smoothing γ -ray spectra at different measurement durations

Эффективность каждого из испытанных методов сильно зависит от времени измерения. При кратковременных измерениях (до 100 с) наиболее эффективным методом (среди испытанных) оказалась скользящая средняя с окном равным пяти каналам. Эта техника увеличивает отношение сигнал / шум на 2–2,5 дБ. Немногим менее эффективна единая для всего энергетического диапазона модель LS-SVM. Использование этих методов теряет смысл при времени измерения более 200–250 с.

В диапазоне времени измерения 10^2-10^3 с наиболее сильное подавление статистического шума удается добиться с помощью экспоненциального сглаживания с $\alpha = 0,75$. Отношение сигнал / шум увеличивается до 2 дБ. Близкую эффективность в этом диапазоне имеет метод скользящей средней с 1 = 3. Он начинает превосходить сглаживающий фильтр с 1 = 5 при времени измерения 80 с и уступает экспоненциальному сглаживанию с $\alpha = 0,75$ при продолжительности измерения более 300 с. Трехоконная регрессия на опорных векторах имеет заметно более низкую эффективность по сравнению с экспоненциальным сглаживанием в рассматриваемом интервале времени измерения.

При времени измерения выше 10³ с полностью исчерпываются возможности шумоподавления с использованием экспоненциального сглаживания, скользящей средней и регрессии на опорных векторах с единой моделью для всего энергетического диапазона. Вместе с тем для времени измерения 1–7·10³ эффективным оказывается шумоподавление с помощью трехоконной регрессии на опорных векторах, оптимизированной для времени измерения 900 с. Поскольку отношение сигнал/шум в необработанном спектре для этого интервала времени измерения составляет 18–22 дБ, эффективность шумоподавления относительно невысокая. Отношение сигнал/шум повышается до 1 дБ.

Заключение

Результаты сравнительного анализа эффективности метода регрессии на несущих векторах (LS-SVM) для подавления статистического шума в спектре γ-излучения, полученном с использованием малогабаритного сцинтилляционного детектора NaI(Tl), показал обоснованность его использования в случае времени измерения 1–7·10³ с. При этом необходима раздельная оптимизация гиперпараметров модели LS-SVM для трех частично перекрывающихся энергетических поддиапазонов с использованием спектрограмм с временем измерения около 900 с.

При времени измерения 7·10³ с LS-SVM ухудшает отношение сигнал/шум. При времени измерения менее 1·10³ с скользящая средняя и экспоненциальный фильтр демонстрируют более высокую эффективность по сравнению с LS-SVM.

Таким образом, можно рекомендовать использование LS-SVM с оптимизацией гиперпараметров модели для трех частично перекрывающихся энергетических поддиапазонов для обработки спектрограмм, полученных с использованием малогабаритного сцинтилляционного детектора NaI(Tl), при времени измерения в пределах 1–7·10³ с.

Библиографические ссылки / References

1. Rouze G, Morgana C, McBratney A. Understanding the utility of aerial gamma radiometrics for mapping soil properties through proximal gamma surveys. *Geoderma*. 2017;289:185–195. DOI: 10.1016/j.geoderma.2016.12.004.

2. Maacha L, Jaffal M, Jarni A. A contribution of airborne magnetic, gamma ray spectrometric data in understanding the structure of the Central Jebilet Hercynian massif and implications for mining. *Journal of African Earth Sciences*. 2017;134:389–403. DOI: 10.1016/j. jafrearsci.2017.07.012.

3. Söderström M, Eriksson J. Gamma-ray spectrometry and geological maps as tools for cadmium risk assessment in arable soils. *Geoderma*. 2013;192:323–334. DOI: 10.1016/j.geoderma.2012.07.014.

4. Khan H, Chaudhry Z, Ismail MK. Assessment of Radionuclides, Trace Metals and Radionuclide Transfer from Soil to Food of Jhangar Valley (Pakistan) Using Gamma-Ray Spectrometry. *Water Air Soil Pollution*. 2010;213:353–362. DOI: 10.1007/s11270-010-0390-4.

5. International atomic energy agency. Guidelines for Radioelement Mapping Using Gamma Ray Spectrometry Data, IAEA-TECDOC-1363. Vienna: IAEA; 2003. 184 p.

6. Blackburn JA. Statistical noise and spectral analysis. *Nuclear Instruments and Methods*. 1968;63(1):66–70. DOI: 10.1016/0029-554X(68)90302-9.

7. Minty B, Hovgaard J. Spectral methods for reducing noise in gamma-ray spectrometry. *ASEG Extended Abstracts*. 2001;4:1–4. DOI: 10.1071/ASEG2001ab089.

8. Varley A, Tyler A, Smith L, Dale P & Davies M. Remediating radium contaminated legacy sites: Advances made through machine learning in routine monitoring of «hot» particles. *Science of The Total Environment*. 2015;521:270–279. DOI: 10.1016/j. scitotenv.2015.03.131.

9. Dragović S, Onjia A. Classification of soil samples according to their geographic origin using gamma-ray spectrometry and principal component analysis. *Journal of Environmental Radioactivity*. 2006;89(2):150–158. DOI: 10.1016/j.jenvrad.2006.05.002.

10. Zhu MH, Liu LG, Qi DX, You Z, Xu AA. Smoothing noisy spectroscopic data with many-knot spline method. *Nuclear Instruments and Methods in Physics Research*. 2008;589:484–486. DOI: 10.1016/j.nima.2008.03.008.

11. Tirosh S, Van De Ville D, Unser M. Polyharmonic smoothing splines and the multidimensional Wiener filtering of fractal-like signals. *IEEE Trans Image Process*. 2006;15 (9):2616–2630. DOI: 10.1109/tip.2006.877390.

12. Claeskens G, Carroll RJ. Automatic estimation of multivariate spectra via smoothing splines. *Biometrika*. 2007;94(2):335–345. DOI: 10.1093/biomet/asm022.

13. Fan CL, Pang MY. Reconstructing smooth curve from noise sampled data. Proceedings. The 2009 International Workshop on Information Security and Application (IWISA 2009). [Place unknown]: Academy Publisher; 2009. p. 218–222.

14. J. B. Lee, A. S. Woodyatt and M. Berman. Enhancement of high spectral resolution remote-sensing data by a noise-adjusted principal components transform. *IEEE Transactions on Geoscience and Remote Sensing*. 1990;28(3):295–304. DOI: 10.1109/36.54356.

15. Hovgaard J. A new processing technique for airborne gamma-ray spectrometer data (Noise adjusted singular value decomposition). Sixth topical meeting on Emergency Preparedness and Response. San Francisco: American Nuclear Society; 1997. p. 123–127.

16. Jiancheng S, Chongxun Z, Yatong Z. Nonlinear Noise Reduction of Chaotic Time Series Based on Multidimensional Recurrent LS-SVM. *Neurocomputing*. 2008;71(16–18):3675–3679. DOI: 10.1016/j.neucom.2008.02.006.

17. Zhang F, O'Donnell LJ. Support vector regression. *Machine Learning. Academic Press.* 2020;12:123–140. DOI: 10.1016/B978-0-12-815739-8.00007-9.

18. Liu J, Guan X, et al. Airborne γ spectrum analysis based on LS-SVM segmented noise reduction method. *Nuclear Electronics and Detection Technology*. 2017;37(1):24–28.

19. Varley A, Tyler A, Dowdall M, Bondar Y, Zabrotski V. An in situ method for the high resolution mapping of ¹³⁷Cs and estimation of vertical depth penetration in a highly contaminated environment. *Science of the Total Environment*. 2017;605–606:957–966. DOI: 10.1016/j.scitotenv.2017.06.067.

20. Kim J, Lim KT, Kim J, Kim C, Jeon B, Park K, Kim G. Quantitative analysis of NaI(Tl) gamma-ray spectrometry using an artificial neural network. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment.* 2019;944:162549. DOI: 10.1016/j.nima.2019.162549.

21. Ying C, Jun L, Lei C. Aeronautical γ spectrum noise reduction method based on LS-SVM segmentation regression. *Journal of Physics*. 2019;2:022141. DOI: 10.1088/1742-6596/1237/2/022141.

Статья поступила в редколлегию 28.03.2023. Received by editorial board 28.03.2023.