

Многопараметрическая калибровка состава эталонных биллонных сплавов по лазерно-искровым эмиссионным спектрам

В. А. Асеев¹⁾, П. С. Колодочки^{2, 3)}, А. А. Савчук¹⁾, М. А. Ходасевич²⁾

¹⁾ Университет ИТМО, Санкт-Петербург, Россия

²⁾ Институт физики им. Б. И. Степанова НАН Беларусь, Минск, Беларусь,
e-mail: m.khodasevich@ifanbel.bas-net.by

³⁾ Университет НАН Беларусь, Минск, Беларусь

Применение методов машинного обучения к лазерно-искровым эмиссионным спектрам эталонных образцов медно-серебряных сплавов позволило получить калибровочные модели их состава с точностью (MAE = 0,36 %, RMSE = 0,45 % и RPD = 18,2), достаточной для нахождения стратиграфических зависимостей состава биллонных артефактов с высоким пространственным разрешением при минимальном повреждении объекта исследования.

Ключевые слова: лазерно-искровая эмиссионная спектроскопия, метод частичных наименьших квадратов, регрессия на опорные вектора.

Multivariate calibration of the composition of reference billon alloys by laser-induced breakdown spectra

V. A. Aseev¹⁾, P. S. Kolodochka^{2, 3)}, A. A. Savchuk¹⁾, M. A. Khodasevich²⁾

¹⁾ ITMO University, Saint-Petersburg, Russia

²⁾ B. I. Stepanov Institute of Physics of the NAS of Belarus, Minsk, Belarus,
e-mail: m.khodasevich@ifanbel.bas-net.by

³⁾ University of the NAS of Belarus, Minsk, Belarus

The application of machine learning methods to laser-induced breakdown spectra of reference samples of copper-silver alloys gives the possibility to create calibration models of their composition with accuracy (MAE = 0.36 %, RMSE = 0.45 % and RPD = 18.2) sufficient to find stratigraphic dependencies of the composition of billon artifacts with high spatial resolution and minimal damage to the object of study.

Keywords: laser-induced breakdown spectroscopy, partial least squares, support vector regression.

При изучении металлических исторических артефактов пристальное внимание уделяется их составу. Сплав драгоценного металла (чаще всего серебра или золота) с недрагоценными металлами, такими как медь, в котором доля драгоценного металла меньше или равна доле недрагоценного металла, называется биллоном [1]. Наиболее распространенным неразрушающим методом научного исследования объектов из биллонных сплавов является рентгено-флуоресцентный анализ (XRF). Методические проблемы, возникающие при применении XRF анализа, связаны с измерением спектров приповерхностного слоя толщиной приблизительно от 30 до 150 мкм и могут существенно искажать результаты. Внутренние слои исторических объектов по составу значительно отличаются от поверхностного из-за воздействия коррозии и поверхностного

обогащения одних элементов за счет других в течение длительного времени нахождения под влиянием окружающей среды [2]. Значительные искажения также может вносить химическая очистка металлических артефактов и неровная поверхность анализируемого объекта. Это приводит к существенным ошибкам при проведении исследований биллонных монет. Такие монеты обладали невысокой покупательной способностью вследствие изготовления из низкопробного серебра и в нынешнее время широко исследуются как археологические артефакты. Для биллонных монет характерна неоднородность состава по поверхности и по глубине, которая могла возникать еще на стадии производства по причине неполного перемешивания компонентов. Результаты XRF анализа поверхности биллонных монет также могут зависеть от примененных методов улучшения внешнего вида. Например, окисление и вытравливание меди с поверхности заготовки из биллона (так называемое отбеливание) [3] или удаление оксидов меди с поверхности с помощью протравки в кислоте после прокаливания и отпуска заготовок приводило к повышению содержания серебра в приповерхностном слое артефактов [4]. Пребывание биллонных монет в земле в течение длительного срока приводит к коррозионным изменениям в их составе. Серебро в них переходит в хлорид серебра, а медь – в оксиды меди. В поверхностном слое оксиды меди могут частично вымываться содержащимися в земле природными кислотами с образованием в артефактах пустот, что увеличивает содержание серебра в нем по сравнению с внутренними слоями [3]. Значительные коррозионные повреждения биллонных монет в почве под воздействием природных факторов в течение нескольких сотен лет обуславливают необходимость проведения реставрационных работ, которые, в свою очередь, также приводят к изменению состава поверхностного слоя. Указанные причины могут приводить к разнице содержания серебра в поверхностном и внутренних слоях низкопробных монет до 10 % [2]. Получение информации о составе внутренних слоев металлических артефактов возможно с помощью применения микрошурфов [1], которые, однако, могут быть загрязнены элементами поверхностного слоя.

Недостатки существующих методов анализа металлических артефактов могут быть преодолены с помощью применения метода лазерно-искровой эмиссионной спектроскопии (LIBS), преимуществами которого являются отсутствие или минимальная пробоподготовка; возможность одновременного определения большого набора химических элементов; проведение анализа с высоким пространственным разрешением; простота и дешевизна, высокая точность калибровки состава артефактов.

Использование методов машинного обучения для анализа многопараметрических эмиссионных спектров при проведении стратиграфического исследования является перспективным. Обычно в этой области применяется однопараметрический подход, использующий нормированные интенсивности на аналитических длинах волн калибруемых элементов или их отношение к интенсивности на реперной длине волны. Например, нормировка на интенсивность линии Ca I в работе [5] позволила при одновременной калибровке до 20 элементов

различать на древнеримских фресках слои аналогичного цвета и распознавать используемую технику рисования. Многопараметрические регрессионные методы являются эталонными, которые в отличие от безэталонного [6] позволяют определять состав объектов исследования с большей точностью, особенно при низких концентрациях неосновных элементов.

На этапе построения многопараметрических калибровочных моделей состава медно-серебряных сплавов были исследованы тринадцать эталонных образцов со следующими концентрациями серебра (в процентах): 72,7; 78,6; 82,9; 83,0; 84,5; 90,5; 92,0; 92,5; 92,7; 97,4; 97,6; 97,7; 98,1. Для каждого эталона зарегистрированы по 200 лазерных эмиссионных спектров в диапазоне длин волн 336,6–892,7 нм с шагом около 0,3 нм (2048 спектральных переменных) в 10 точках на поверхности. Предобработка спектров заключалась в проведении коррекции базовой линии с помощью аддитивного метода наименьших квадратов со штрафом и итеративным пересчетом весов (airPLS – adaptive iteratively reweighted Penalized Least Squares) [7]. Преимуществом метода airPLS является то, что для него не требуется предварительная информация об обнаруженных пиках и особенностях спектра. Для уменьшения влияния выбросов на точность калибровочных моделей все спектры, суммарная интенсивность которых выходит за пределы $\pm 10\%$ от средней величины по всем импульсам для каждого эталона, изымались из рассмотрения.

При построении калибровочных моделей линейным методом частичных наименьших квадратов (partial least squares – PLS) [8] и нелинейной регрессией на опорные вектора (support vector regression – SVR) [9] входные данные (спектры и величины концентраций серебра, соответственно) делятся на обучающую и проверочную выборки. Обучающая выборка содержит приблизительно 60 % от количества образцов, оставшиеся составляют проверочную. В рассматриваемом случае обучение проводилось по 7 образцам, проверка – по 6. В общем случае, при большом количестве измерений предпочтителен метод случайного разделения данных (random sampling – RS) [10]. Недостатком этого метода является плохая обобщаемость моделей, которая устраняется статистическим усреднением нескольких моделей. Применяемая нами модификация метода RS заключается в обязательном выборе образцов с двумя минимальными и двумя максимальными концентрациями калибруемого параметра в обучение, трехкратном повторении построения калибровочных моделей с последующим их усреднением без деления на обучающую и проверочную выборку. Все рассматриваемые многопараметрические модели являются интерполяционными и характеризуются более высокой способностью обобщения по сравнению с одиночной моделью, построенной на выборке, которая создана методом RS.

Показателями качества многопараметрических моделей являются среднее абсолютное отклонение (mean absolute error – MAE), среднеквадратичное отклонение (root mean square error – RMSE) и остаточное отклонение (residual prediction deviation – RPD), представляющее собой отношение стандартного отклонения концентрации серебра в рассматриваемой выборке к RMSE.

На первом этапе проведения калибровки строится широкополосная модель PLS, по минимуму RMSE в проверочной выборке которой выбирается оптимальное количество латентных переменных. По широкополосной модели SVR определяется оптимальная функция преобразования ядра. Затем для выбора спектральных переменных применяется оригинальная модификация [11] метода поиска комбинации движущихся окон (search combination moving window interval – scmwi) для уменьшения мультиколлинеарности и избыточности данных. При построении моделей scmwiPLS и scmwiSVR ширина спектрального окна превышает оптимальное количество латентных переменных для PLS на 1. На рис. 1 представлены калибровочные зависимости концентрации серебра в исследуемых эталонных сплавах с медью, построенные линейным методом PLS и нелинейным методом SVR без и с выбором спектральных переменных поиском комбинации движущихся окон оптимальной ширины. В таблице приведены показатели качества созданных калибровочных моделей.

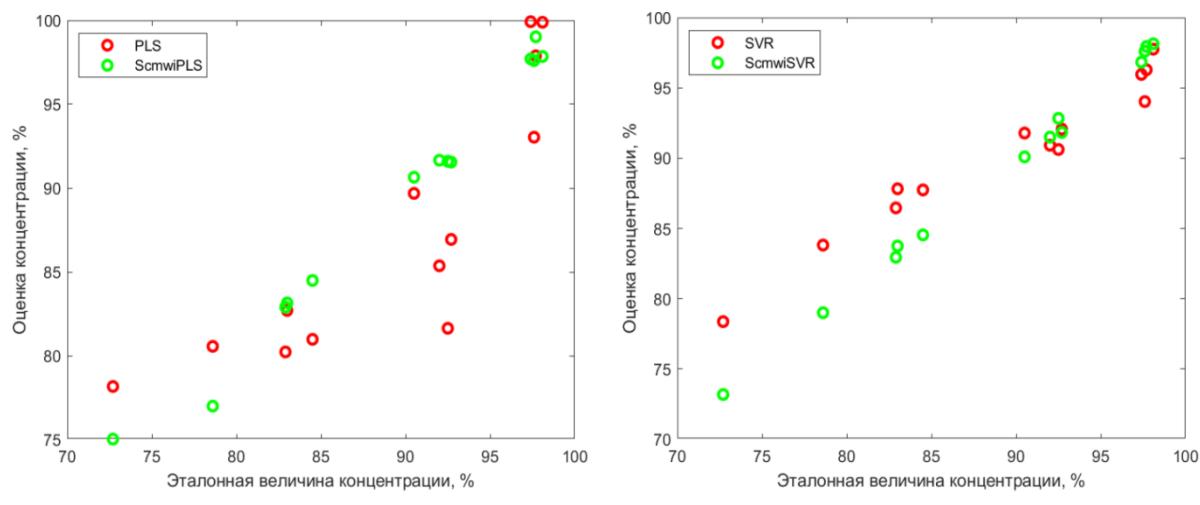


Рис. 1. Калибровочные зависимости концентрации серебра в исследуемых эталонных сплавах, построенные методами:
а – методом PLS и scmwiPLS; б – методами SVR и scmwiSVR

Среднее абсолютное, среднеквадратичное и остаточное отклонение калибровок концентрации серебра в эталонных сплавах по эмиссионным лазерным спектрам в диапазоне длин волн 336,6-892,7 нм с шагом около 0,3 нм без и с выбором спектральных переменных методом поиска комбинации движущихся окон оптимальной ширины

	PLS	SVR	scmwiPLS	scmwiSVR
MAE, %	3,62	2,63	0,65	0,36
RMSE, %	4,64	3,16	0,97	0,45
RPD	1,8	2,6	8,5	18,2

Из приведенных результатов видно, что модели без выбора спектральных переменных являются качественными ($RPD < 3$), а с выбором спектральных

переменных – количественными ($RPD > 3$). Лучшие показатели точности калибровки ($MAE = 0,36\%$, $RMSE = 0,45\%$ и $RPD = 18,2$) достигаются усредненной нелинейной многопараметрической регрессией на опорные вектора с поиском комбинации движущихся спектральных окон оптимальной ширины при формировании обучающих выборок модифицированным методом случайного отбора. Достигнутая точность определения концентрации серебра является достаточной для нахождения стратиграфических зависимостей состава артефактов с высоким пространственным разрешением при минимальном повреждении объекта исследования и позволит получить объективную информацию о металлических артефактах.

Работа выполнена при финансовой поддержке Белорусского фонда фундаментальных исследований (проект Ф25МС-005).

Библиографические ссылки

1. Горлов К. В. «... И начаша пенязми торговати»: О находках ливонских монет конца XIV–начала XV века в Пскове (по материалам мстиславских раскопов) / К. В. Горлов, Е. В. Салмина, С. А. Салмин // Caurus. 2022. №. 1. С. 11–25.
2. Ениосова Н. В., Митоян Р. А. Рентгеноспектральный метод анализа археологического металла: преимущества, ограничения и ловушки в процессе измерения и интерпретации // Труды IV (XX) Всероссийского археологического съезда в Казани. 2014. С. 143–146.
3. Пономарев А. Л. Эволюция денежных систем Причерноморья и Балкан в XIII–XV вв. М.: МГУ, 2012. 670 с.
4. Морозов А. А. Монетное дело в Крымском ханстве при Мухаммеде I Гирее и денежная реформа 922 г.х. // Расмир: Восточная нумизматика. Вторая Международная научная конференция. 2015. С. 175–215.
5. Stratigraphic characterization of ancient Roman frescos by laser induced breakdown spectroscopy and importance of a proper choice of the normalizing lines / V. Lazić [et al.] // Spectrochimica Acta Part B: Atomic Spectroscopy. 2020. Vol. 168. Art ID 105853.
6. Gaudiuso R. Calibration-free inverse method for depth-profile analysis with laser-induced breakdown spectroscopy // Spectrochimica Acta Part B: Atomic Spectroscopy. 2016. Vol. 123. P. 105–113.
7. Zhang Z. M. Baseline Correction Using Adaptive Iteratively Reweighted Penalized Least Squares / Z. M. Zhang, S. Chen, Y. Z. Liang // Analyst. 2010. Vol. 135. P. 1138–1146.
8. Garson G. D. Partial Least Squares (PLS-SEM) // Statistical Publishing Associates, 2016.
9. Guenther N., Schonlau M. Support Vector Machines // The Stata Journal: Promoting Communications on Statistics and Stata. 2016. Vol. 16, No. 4. P. 917–937.
10. Genç B., Tunç H. Optimal training and test sets design for machine learning // Turkish Journal of Electrical Engineering and Computer Sciences. 2019. Vol. 27. ArtID. 60.
11. Ходасевич М. А., Саскевич Н. А. Методы построения обучающего набора для калибровки с помощью флуоресцентной спектроскопии небольших выборок // Весці НАН Беларусі, сер. фіз.-мат. науک. 2018. № 1. С. 77–83.