

## МНОГОПАРАМЕТРИЧЕСКАЯ КАЛИБРОВКА СОСТАВА НИЗКОЛЕГИРОВАННЫХ СТАЛЕЙ ПО ПРЕДОБРАБОТАННЫМ ЭМИССИОННЫМ СПЕКТРАМ НИЗКОГО РАЗРЕШЕНИЯ С ВЫБОРОМ СПЕКТРАЛЬНЫХ ПЕРЕМЕННЫХ

М. В. Бельков, К. Ю. Кацалап, Д. А. Королько\*, М. А. Ходасевич

УДК 543.423

<https://doi.org/10.47612/0514-7506-2023-90-2-174-179>

Институт физики НАН Беларуси,  
Минск, Беларусь; e-mail: d.korolko@ifanbel.bas-net.by

(Поступила 24 января 2023)

По эмиссионным спектрам низкого разрешения лазерно-искровой плазмы эталонных образцов низколегированных сталей осуществлена калибровка концентраций C, Mn, Si, Cr, Ni и Cu. На основе предварительной обработки данных в виде нормировки спектров на длине волны эмиссионной линии Fe II 252.0609 нм и вычета базовой линии, а также выбора спектральных переменных оригинальным методом поиска комбинации движущихся окон для метода частичных наименьших квадратов построены многопараметрические калибровочные модели для указанных элементов со следующими характеристиками: для C (в диапазоне концентраций до 0.7 %) среднеквадратичное и остаточное отклонения в проверочной выборке составили 0.04 % и 4.7, Mn (до 1.9 %) 0.02 % и 24.8, Si (до 0.9 %) 0.01 % и 12.9, Cr (до 1 %) 0.01 % и 21.8, Ni (до 0.7 %) 0.007 % и 23.3, Cu (до 0.5 %) 0.006 % и 23.2 соответственно. Модели являются количественными (остаточное отклонение > 3) для шести рассматриваемых элементов, включая C.

**Ключевые слова:** лазерно-искровая эмиссионная спектроскопия, метод частичных наименьших квадратов, выбор спектральных переменных, низколегированные стали.

Calibration of concentrations of C, Mn, Si, Cr, Ni, and Cu by the low-resolution laser induced breakdown spectroscopy is made in reference to low-alloy steels etalons. Data preprocessing in the form of spectrum normalization at Fe II 252.0609 nm emission line wavelength and baseline correction, as well as spectral variables selection with an original method of searching combination moving window for the partial least squares method made it possible to build multivariate calibration models for all considered elements with the following characteristics: for C (in the concentration range up to 0.7%) the value of root-mean-square error and residual predictive deviation is 0.04 % and 4.7, Mn (up to 1.9%) – 0.02 % and 24.8, Si (up to 0.9%) – 0.01 % and 12.9, Cr (up to 1%) – 0.01 % and 21.8, Ni (up to 0.7%) – 0.007 % and 23.3, Cu (up to 0.5%) – 0.006 % and 23.2, respectively. Models are quantitative (residual predictive deviation > 3) for all six elements considered, including carbon.

**Keywords:** laser induced breakdown spectroscopy, partial least squares, spectral variable selection, low-alloy steels.

**Введение.** Сталь и другие сплавы на основе железа являются базовыми материалами, используемыми в строительстве различных конструкций, автомобилей, кораблей, машин, инструментов и бытовой техники. Особые ее свойства — ударная вязкость, прочность, твердость, коррозионная стойкость — определяются в основном технологическими примесями и легирующими добавками, которые можно разделить на две категории: неметаллические (углерод, азот, фосфор, сера) и металличе-

## MULTIVARIATE CALIBRATION OF THE COMPOSITION OF LOW-ALLOY STEELS USING PRE-PROCESSED LOW-RESOLUTION EMISSION SPECTRA WITH SPECTRAL VARIABLES SELECTION

M. V. Belkov, K. Y. Catsalap, D. A. Korolko\*, M. A. Khodasevich (B. I. Stepanov Institute of Physics of the National Academy of Sciences of Belarus, Minsk, Belarus; e-mail: d.korolko@ifanbel.bas-net.by)

ские (марганец, хром, никель). Марганец является эффективным раскислителем и десульфуризатором и может улучшить закалочную способность и горячую обрабатываемость стали. Хром позволяет значительно повысить прочность, стойкость к окислению, твердость, коррозионную износостойкость стали. Поэтому для классификации или сортировки важно точно определить концентрацию неосновных элементов в стали. Для сталелитейной промышленности актуальны технологический анализ расплавленных сталей, анализ стальных материалов в полевых условиях, сортировка стальных заготовок или отходов в процессе производства. Для определения состава сталей обычно применяют методы масс-спектрометрии [1], оптической эмиссионной спектроскопии с использованием искрового разряда [2], индуктивно-связанной плазмы [3] или тлеющего разряда [4], лазерно-искровой эмиссионной спектроскопии (LIBS) [5].

Метод LIBS на сегодняшний день считается полуколичественным [6], но в то же время имеет ряд преимуществ: возможность экспрессного анализа материалов на открытом воздухе, относительно низкая стоимость и отсутствие необходимости предварительной обработки образцов. Такие достоинства могут стать определяющими, особенно в управлении промышленными процессами, где применение других аналитических методов связано с серьезными трудностями. Недостатки LIBS-анализа сталей и их сплавов, связанные с нестационарностью лазерно-искровой плазмы, ведущей к нестабильности перекрывающихся эмиссионных линий железа с линиями исследуемых примесей и добавок, и флуктуациями интенсивности линий и непрерывного излучения, не позволяют разработать калибровочные модели высокого качества с помощью методов классической спектроскопии с низким разрешением в компактных устройствах. В этом случае большим потенциалом обладают многопараметрические модели, учитывающие большое количество спектральных переменных. Поэтому разработка более совершенных многопараметрических методов обработки эмиссионных спектров имеет первостепенное значение для применения LIBS в сортировке и анализе сталей.

**Эксперимент.** Исследованы 65 эталонных образцов стали: УГ0д-УГ7д, УГ9д, УГ17е-УГ21е, УГ75, УГ79, УГ82, УГ83, УГ84, УГ86, УГ88-УГ97 производства ИСО, Россия; 51/1-58/1, 63/2-65/2, 72-76, 101-103, 110-125 производства IMZ, Польша, легированных примесями следующих химических элементов: С (в диапазоне не более 0.8 %), Мn (2.0 %), Si (1.2 %), Cr (1.0 %), Ni (0.8 %) и Cu (0.5 %). Для возбуждения лазерно-эрозионной плазмы использован двухканальный YAG:Nd-лазер [7]. Для каждого образца с помощью малогабаритного спектрометра FireFly 4000 зарегистрированы по 50 эмиссионных спектров в диапазоне 172—507 нм с разрешением 0.4 нм и шагом 0.1 нм в пяти точках на поверхности. Экспериментальная установка и условия измерений описаны в [8].

**Результаты и их обсуждение.** В предыдущей работе [9] для калибровки концентрации примесей в сталях был использован широкополосный метод частичных наименьших квадратов (PLS) [10] — билинейный статистический метод, который одновременно проецирует матрицу спектров и столбец концентрации в новое маломерное пространство латентных переменных, характеризуемое условием максимальной корреляции проекций спектров и проекции концентрации в этом пространстве. Метод дает хорошие результаты в случае, когда матрица спектров мультиколлинеарна и количество образцов значительно меньше числа спектральных переменных. Для проведения калибровки матрица спектров и вектор концентрации должны быть разделены на обучающий и проверочный наборы данных. В [9] для этой цели был применен алгоритм Кеннарда—Стоуна [11]. Образец с концентрацией, наиболее близкой к середине рассматриваемого диапазона, выбирается первым в набор обучающих данных. Остальными образцами обучающего набора являются наиболее удаленные по концентрации от уже выбранных. Количество образцов в обучающей и проверочной выборках должно примерно соответствовать соотношению 3:2 [12]. Одним из критериев качества многопараметрических моделей является среднеквадратичное отклонение калибровки RMSEP в образцах проверочной выборки. Построенные калибровочные модели по всему измеренному диапазону длин волн характеризуются следующими значениями RMSEP: 0.06 % для С, 0.12 % для Мn, 0.09 % для Si, 0.13 % для Cr, 0.07 % для Ni и 0.08 % для Cu, сравнение с которыми позволит определить целесообразность применения методов предобработки спектров, формирования выборок образцов и выбора спектральных переменных.

В данной работе до построения калибровочных моделей найдены и исключены выбросы в серии последовательных спектров, измеренных в каждой из пяти точек на поверхности образца. Критерий выброса — превышение суммарной интенсивностью спектра 10 % отклонения от средней величины по всем используемым для построения модели лазерным импульсам. После выявления выбросов проведено усреднение оставшихся спектров по всем пяти точкам. В отличие от моделей [8, 9] перед нормировкой спектров на интенсивность выбранной эмиссионной линии Fe II 252.0609 нм в качестве

предварительной обработки данных применена коррекция базовой линии. Среди большого количества разработанных методов коррекции базовой линии спектров [13] чаще всего используется один из видов полиномиальной аппроксимации. Недостатками таких методов являются интерактивность и невысокая точность в условиях низкого отношения сигнал/шум.

Используемый адаптивный метод наименьших квадратов со штрафом и итеративным пересчетом весов (airPLS) [14] не требует какого-либо интерактивного вмешательства и предварительного обнаружения особенностей в спектре. Целевой функцией airPLS является

$$Q = \sum_{i=1}^{m=3630} w_i |x_i - z_i|^2 + \lambda \sum_{j=2}^{m=3630} |z_j - z_{j-1}|^2,$$

где первое слагаемое — сумма взвешенных коэффициентами  $w_i$  разностей экспериментальных значений спектральных переменных  $x_i$  и отсчетов базовой линии  $z_i$ , второе слагаемое характеризует гладкость базовой линии,  $m$  — количество спектральных переменных. Единственным задаваемым коэффициентом  $\lambda = 2 \cdot 10^7$  определяется компромисс между точностью аппроксимации и гладкостью базовой линии. В начале работы алгоритма коэффициенты  $w_i$  принимаются равными единице. Каждая последующая итерация нахождения базовой линии использует разницу между значениями спектральных переменных и отсчетами базовой линии, определенными на предыдущем этапе. Если разница больше нуля, то соответствующие коэффициенты обнуляются, если меньше, то  $w_i^t = \exp[t(x_i - z_i^{t-1})/|\mathbf{d}^t|]$ , где  $t$  — номер итерации, вектор  $\mathbf{d}$  состоит из отрицательных элементов разности векторов  $\mathbf{x}$  и  $\mathbf{z}^{t-1}$ . Реализация алгоритма airPLS доступна в виде программного обеспечения с открытым исходным кодом [15]. Версия для пакета MATLAB реализована на основе разреженных матриц и требует исключительно малых вычислительных ресурсов.

Для расширения диапазонов концентраций проверочных выборок, среднеквадратичное отклонение калибровки по которым является целевой функцией применения метода частичных наименьших квадратов, обучающая и проверочная выборки формировались случайным образом. При этом выполнялось условие обязательного отбора в обучающую выборку двух эталонных образцов с минимальными калибруемыми концентрациями и двух — с максимальными. После построения калибровочной зависимости две пробы с экстремальными значениями концентрации были исключены из рассмотрения, что незначительно сузило диапазон моделирования. При использовании такого метода калибровочная модель является интерполяционной и ее точность менее подвержена влиянию краевых по концентрации образцов. По причине случайного характера формирования выборок моделирование повторялось 10 раз для каждого элемента. Окончательный отбор модели проводился по худшей сумме величин среднеквадратичного отклонения калибруемой концентрации в обеих выборках, что гарантирует лучшие показатели моделей при ином выборе образцов в выборках.

Особенностью разработанных многопараметрических моделей является также применение выбора спектральных переменных. Многомерный анализ данных оперирует латентными переменными, которые содержат только важную для моделирования информацию, извлеченную из мультиколлинеарных широкополосных спектров. Классический одно- или маломерный подход использует имеющие четкий физический смысл спектральные переменные. Классический и многопараметрический подходы принципиально различаются в плане использования спектральной информации. Наилучшие результаты можно получить при компромиссном подходе с выбором спектральных переменных, который позволяет предотвратить переоценку многопараметрических моделей [16]. Методы выбора спектральных переменных эффективны при работе с большими объемами сильно коррелированных данных и нацелены на определение подмножества переменных, описывающих большую часть важной для построения многопараметрической модели информации. Известно большое количество подходов, которые использовались для извлечения полезной информации в различных областях спектроскопии. Среди рассмотренных нами ранее методов выбора спектральных переменных (интервальные методы PLS, генетический алгоритм, ранжирование спектральных переменных по их коэффициенту корреляции с искомым параметром [17], алгоритм последовательного проецирования [18]) выбрана оригинальная модификация [19] поиска комбинации движущихся спектральных окон [20] в интервальном методе PLS (scmwiPLS).

Метод scmwiPLS использует не отдельные спектральные переменные, а спектральные интервалы — окна. На первом этапе с помощью широкополосной модели определяется оптимальное количество латентных переменных по минимальному значению среднеквадратичного отклонения калибровки по проверочной выборке. На втором этапе ширина окна фиксируется на одну спектральную переменную больше, чем оптимальное количество латентных переменных в широкополосной модели. Это усло-

вие минимизирует ширину окна, сохраняя при этом возможность выбора латентных переменных даже в одном окне. Первое окно сдвигается на одну спектральную переменную за шаг и разрабатывается многопараметрическая модель с использованием метода PLS. После прохождения первым окном всего диапазона, оптимальное положение этого окна определяется минимальным среднеквадратичным отклонением и фиксируется. Процедура повторяется для последующих спектральных окон до включения в модель всего спектрального диапазона. В этом случае моделирование осуществляется путем объединения спектральных переменных, принадлежащих уже фиксированным окнам и сдвигающемуся окну. На третьем этапе выбирается модель с минимальной величиной RMSEP в зависимости от количества окон. На рис. 1 представлен эмиссионный спектр эталона низколегированной стали и выбранные методом scmwiPLS спектральные переменные для калибровки Mn.

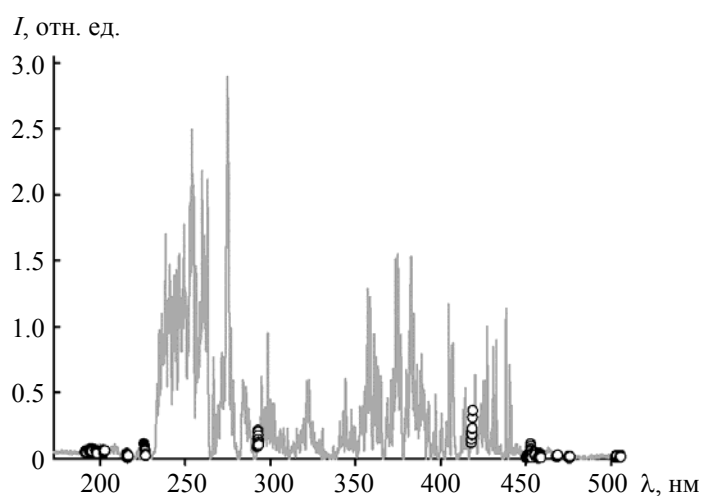


Рис. 1. Эмиссионный спектр эталона низколегированной стали и выбранные методом scmwiPLS спектральные переменные для калибровки Mn (точки)

Результаты калибровки марганца приведены в табл. 1, где представлены количество выбранных спектральных переменных, среднеквадратичное отклонение в обучающих и проверочных выборках образцов и остаточное отклонение предсказания (RPD), равное отношению стандартного отклонения калибруемого параметра в выборке к среднеквадратичному отклонению калибровки. Чем выше RPD, тем ближе калибровочная зависимость к прямой. Пороговое значение RPD, при котором калибровка считается количественной, равно трем [21], поэтому все представленные модели являются количественными.

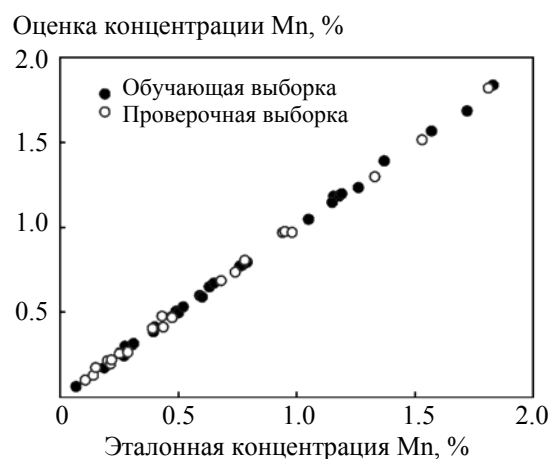
**Т а б л и ц а 1. Результаты калибровки марганца для 10 вариантов наборов обучающей и проверочной выборок**

№	Количество выбранных спектральных переменных	RMSE обучающей выборки, %	RPD обучающей выборки	RMSE проверочной выборки, %	RPD проверочной выборки
1	286	0.03	19.1	0.01	36.2
2	252	0.02	26.7	0.02	32.9
3	325	0.02	31.4	0.01	34.5
4	198	0.02	24	0.02	21.8
5	300	0.02	33.6	0.01	30.9
6	136	0.008	63.7	0.02	23.0
7	319	0.01	41.2	0.02	20.9
8	176	0.02	25.4	0.02	24.8
9	310	0.02	32.6	0.03	15.1
10	336	0.02	30.3	0.02	19.1

Для дальнейшего построения моделей для каждого элемента выбран худший по точности вариант с близкими RPD в обучающей и проверочной выборках из 10 случайных наборов (№ 8 для марганца). Характеристики выбранных моделей для всех калибруемых концентраций показаны в табл. 2. Все шесть калибровочных моделей являются количественными. Можно сделать вывод о сбалансированности моделей, так как параметры калибровки по всем элементам близки в обоих наборах данных. Соответствие эталонной концентрации марганца и оценки, полученной с помощью разработанной многопараметрической калибровочной модели с выбором спектральных переменных, представлено на рис. 2.

**Т а б л и ц а 2.** Характеристики калибровочных моделей для исследуемых элементов

Элемент	RMSE обучающей выборки, %	RPD обучающей выборки	Концентрация для обучающей выборки, %	RMSE проверочной выборки, %	RPD проверочной выборки	Концентрация для проверочной выборки, %
C	0.05	4.5	0.003—0.750	0.04	4.7	0.029—0.640
Mn	0.02	25.4	0.008—1.970	0.02	24.8	0.106—1.810
Si	0.02	11.4	0.006—1.060	0.01	12.9	0.043—0.790
Cr	0.01	17.0	0.008—0.998	0.01	21.8	0.035—0.940
Ni	0.008	20.9	0.005—0.690	0.007	23.3	0.039—0.600
Cu	0.007	18.8	0.004—0.620	0.006	23.2	0.012—0.460



**Рис. 2.** Соотношение между предсказанной с помощью smwiPLS модели и эталонной концентрацией Mn

Таким образом, по сравнению с [9] при использовании одних и тех же эмиссионных спектров эталонных образцов низколегированных сталей за счет применения коррекции базовой линии и иного метода формирования обучающей и проверочной выборок получены калибровочные модели с лучшими показателями среднеквадратичного отклонения и расширенным диапазоном концентраций шести элементов. Для C диапазон расширился от 0.130—0.430 до 0.029—0.640, для Mn от 0.470—1.150 до 0.106—1.810, для Si от 0.150—0.330 до 0.043—0.790, для Cr от 0.090—0.430 до 0.035—0.940, для Ni от 0.050—0.250 до 0.039—0.600 и для Cu от 0.060—0.260 до 0.012—0.460.

**Заключение.** Предварительная обработка спектров с помощью адаптивного метода наименьших квадратов со штрафом и итеративным пересчетом весов и нормировки на интенсивность эмиссионной линии железа Fe II 252.0609 нм и выбор спектральных переменных с помощью поиска комбинации движущихся окон в методе частичных наименьших квадратов повышают качество многопараметрических моделей калибровки концентраций технологических примесей и добавок в низколегированных сталях по спектрам, полученным с помощью лазерно-искровой эмиссионной спектроскопии низкого разрешения, и делают калибровку количественной для рассматриваемых элементов. Для калибровочной модели углерода среднеквадратичное отклонение  $RMSEP = 0.04 \%$ , остаточное

отклонение в проверочной выборке  $RPDP = 4.7$ , для Mn — 0.02 % и 24.8, Si — 0.01 % и 12.9, Cr — 0.01 % и 21.8, Ni — 0.007 % и 23.3, Cu — 0.006 % и 23.2.

- [1] Y. Wei, R. S. Varanasi, T. Schwarz, L. Gomell, H. Zhao, D. J. Larson, B. Sun, G. Liu, H. Chen, D. Raabe, B. Gault. *Patterns*, **2** (2021) 100192
- [2] S. Grunberger, S. Eschbock-Fuchs, J. Hofstadler, A. Pissenberger, H. Duchaczek, S. Trautner, J. D. Pedarnig. *Spectrochim. Acta B: At. Spectrosc.*, **169** (2020) 105884
- [3] M. W. Vaughan, P. Samimi, S. L. Gibbons, R. A. Abrahams, R. C. Harris, R. E. Barber, I. Karaman. *Scripta Mater.*, **184** (2020) 63—69
- [4] Ch. J. Rao, S. Ningshen, J. Philip. *Spectrochim. Acta B: At. Spectrosc.*, **172** (2020) 105973
- [5] L. Quackatz, A. Griesche, T. Kannengiesser. *Forces in Mechanics*, **6** (2022) 100063
- [6] V. Motto-Ros, D. Syvilay, L. Bassel, E. Negre, F. Trichard, F. Pelascini, J. El Haddad, A. Harhira, S. Moncayo, J. Picard, D. Devismes, B. Bousquet. *Spectrochim. Acta B: At. Spectrosc.*, **140** (2018) 54—64
- [7] T. A. Labutin, A. M. Popova, S. M. Zaytsev, N. B. Zorova, M. V. Belkov, V. V. Kiris, S. N. Raikov. *Spectrochim. Acta B*, **99** (2014) 94—100
- [8] М. В. Бельков, Д. А. Борисевич, К. Ю. Кацалап, М. А. Ходасевич. *Журн. прикл. спектр.*, **88**, № 5 (2021) 737—742 [M. V. Belkov, D. A. Borisevich, K. Y. Catsalap, M. A. Khodasevich. *J. Appl. Spectr.*, **88** (2021) 970—974]
- [9] M. V. Belkov, D. A. Borisevich, K. Yu. Katsalap, M. A. Khodasevich. *Opt. and Spectr.*, **10** (2022) 1611—1616
- [10] P. Geladi, B. R. Kowalski. *Anal. Chim. Acta*, **185** (1986) 1—17
- [11] S. Nawar, A. M. Mouazen. *Comput. Electron. Agric.*, **151** (2018) 469—477
- [12] Standard Practices for Infrared Multivariate Quantitative Analysis (СТ РК АСТМ Е 1655-2011 “Стандартные методы по инфракрасному многомерному количественному анализу”)
- [13] G. Schulze, A. Jirasek, M. M. L. Yu, A. Lim, R. F. B. Turner, M. W. Blades. *Appl. Spectrosc.*, **59** (2005) 545—574
- [14] Z. M. Zhang, S. Chen, Y. Z. Liang. *Analyst*, **135** (2010) 1138—1146
- [15] <https://code.google.com/archive/p/airpls/>
- [16] A. A. Gomes, S. M. Azcarate, P. H. Gonçalves Dias Diniz, D. D. S. Fernandes, G. Veras. *Food Chem.*, **370** (2022) 1—13
- [17] Z. Xiaobo, Z. Jiewen, M. J. W. Povey, M. Holmes, M. Hanpin. *Anal. Chim. Acta*, **667** (2010) 14—32
- [18] S. F. C. Soares, A. A. Gomes, M. C. U. Araujo, A. R. G. Filho, R. K. H. Galvão. *Trends in Anal. Chem.*, **42** (2013) 84—98
- [19] M. Khodasevich, V. Aseev. *Opt. and Spectrosc.*, **124** (2018) 748—752
- [20] Y. P. Du, Y. Z. Liang, J. H. Jiang, R. J. Berry, Y. Ozaki. *Anal. Chim. Acta*, **501** (2004) 183—191
- [21] R. Zornoza, C. Guerrero, J. Mataix-Solera, K. M. Scow, V. Arcenegui, J. Mataix-Beneyto. *Soil Biol. and Biochem.*, **40** (2008) 1923—1930