

УДК УДК 535.243+663.2

Применение многопараметрического анализа широкополосных спектров пропускания для калибровки физико-химических показателей вин

М.А. Ходасевич¹, Е.А. Скорбанова², М.В. Роговая¹

¹Институт физики Национальной академии наук Беларуси,
пр. Независимости, 68, г. Минск 220072, Беларусь

²Научно-практический институт плодоводства, виноградарства и пищевых технологий,
ул. Виерул, 59, муниц. Кишинев, п.г.т. Кодру 2070, Молдова

Поступила 12.02.2019

Принята к печати 23.05.2019

Использованию многопараметрической обработки спектральной информации в последнее время отдается предпочтение в связи с экспрессным характером этого метода, простотой применения математических пакетов, отсутствием необходимости введения дополнительных реагентов. Целью работы являлось применение методов многопараметрического анализа широкополосных спектров пропускания для калибровки физико-химических показателей вин и повышение точности этой калибровки с помощью выбора спектральных переменных.

На примере некупажированных сортовых молдавских вин показано, что применение метода интервальной проекции на латентные структуры по комбинации движущихся окон в спектрах пропускания вин в диапазоне 220–2500 нм позволяет существенно уменьшить среднеквадратичную ошибку калибровки по сравнению с широкополосными многопараметрическими методами. Величины остаточного отклонения предсказания, превышающие пороговое значение 2,5 для содержания К, Са, Mg, щавелевой, яблочной и янтарной кислот, 2,3-бутиленгликоля, золы и фенольных соединений для красных вин и Mg, винной, лимонной и молочной кислот, 2,3-бутиленгликоля, золы, фенольных соединений и растворимых солей для белых вин, демонстрируют хорошее качество калибровки.

Применение предложенного метода калибровки физико-химических параметров вин позволяет заменить традиционные методы на проведение спектральных измерений, доступное не только в лабораторных, но и в полевых условиях, и характеризующееся малыми величинами среднеквадратичной ошибки калибровки.

Ключевые слова: многопараметрический анализ, спектр пропускания, интервальная проекция на латентные структуры.

DOI: 10.21122/2220-9506-2019-10-2-198-206

Адрес для переписки:

М.А. Ходасевич
Институт физики НАН Беларуси,
пр. Независимости, 68, г. Минск 220072, Беларусь
e-mail: m.khodasevich@ifanbel.bas-net.by

Address for correspondence:

M.A. Khodasevich
Institute of Physics of the National Academy of Science of Belarus,
Nezavisimosti Ave., 68, Minsk 220072, Belarus
e-mail: m.khodasevich@ifanbel.bas-net.by

Для цитирования:

М.А. Ходасевич, Е.А. Скорбанова, М.В. Роговая.
Применение многопараметрического анализа широкополосных спектров пропускания для калибровки физико-химических показателей вин.
2019. – Т. 10, № 2. – С. 198–206.
DOI: 10.21122/2220-9506-2019-10-2-198-206

For citation:

M.A. Khodasevich, E.A. Scorbanov, M.V. Rogovaya.
[Application of Multivariate Analysis of Broadband Transmission Spectra for Calibration of Physico-Chemical Parameters of Wines].
Devices and Methods of Measurements.
2019, vol. 10, no. 2, pp. 198–206 (in Russian).
DOI: 10.21122/2220-9506-2019-10-2-198-206

Application of Multivariate Analysis of Broadband Transmission Spectra for Calibration of Physico-Chemical Parameters of Wines

M.A. Khodasevich¹, E.A. Scorbanov², M.V. Rogovaya¹

¹*Institute of Physics of the National Academy of Science of Belarus,
Nezavisimosti Ave., 68, Minsk 220072, Belarus*

²*Scientific and Practical Institute of Horticulture, Viticulture and Food Technologies,
Kishinev, Moldova*

Received 12.02.2019

Accepted for publication 23.05.2019

Abstract

The use of multivariate processing of spectral information has recently been favored due to the express nature of this method, the ease of use of mathematical packages, and the lack of the need to add chemical reagents. The aim of the work is using the methods of multivariate analysis of broadband transmission spectra to calibrate the physicochemical parameters of wines and to improve the accuracy of this calibration by selecting spectral variables.

Using the interval projection to latent structures of the transmission spectra in the range of 220–2500 nm, the physicochemical characteristics of the varietal unblended Moldovan wine are calibrated. Interval methods of multivariate data analysis allow significantly reducing the root mean square calibration error in comparison with the broadband multivariate methods. Residual predictive deviations exceed the threshold value of 2.5 for K, Ca, Mg, oxalic, malic and succinic acids, 2,3-butylene glycol, ash and phenolic compounds for red wines and Mg, tartaric, citric and lactic acids, 2,3-butylene glycol, ash, phenolic compounds and soluble salts for white wines. These values demonstrate good calibration quality.

The application of the proposed method for calibrating the physicochemical parameters of wines makes it possible to replace traditional methods with spectral measurements, which are available not only in laboratory but also in the field, and characterized by small values of the root mean square error of calibration.

Keywords: multivariate analysis, transmission spectrum, interval projection to latent structures.

DOI: 10.21122/2220-9506-2019-10-2-198-206

Адрес для переписки:

М.А. Ходасевич
Институт физики НАН Беларуси,
пр. Независимости, 68, г. Минск 220072, Беларусь
e-mail: m.khodasevich@ifanbel.bas-net.by

Address for correspondence:

M.A. Khodasevich
Institute of Physics of the National Academy of Science of Belarus,
Nezavisimosti Ave., 68, Minsk 220072, Belarus
e-mail: m.khodasevich@ifanbel.bas-net.by

Для цитирования:

М.А. Ходасевич, Е.А. Скорбанова, М.В. Роговая.
Применение многопараметрического анализа широкополосных спектров пропускания для калибровки физико-химических показателей вин.
2019. – Т. 10, № 2. – С. 198–206
DOI: 10.21122/2220-9506-2019-10-2-198-206

For citation:

M.A. Khodasevich, E.A. Scorbanov, M.V. Rogovaya.
[Application of Multivariate Analysis of Broadband Transmission Spectra for Calibration of Physico-Chemical Parameters of Wines].
Devices and Methods of Measurements.
2019, vol. 10, no. 2, pp. 198–206 (in Russian).
DOI: 10.21122/2220-9506-2019-10-2-198-206

Введение

Виноградные вина являются сложными смесями химических соединений. Их состав зависит от сорта винограда, почвенно-климатических особенностей географической зоны его произрастания, использования регламентированных агротехнических приемов и других факторов. Кроме этилового спирта, содержание которого в сухих виноградных винах может достигать до 15 % об., в вине содержатся в разных количествах высшие спирты, летучие и нелетучие кислоты, сложные эфиры, многоатомные спирты, щелочные металлы К, Na, Ca, Mg и многие другие компоненты. Определение этих параметров осуществляется с использованием современных инструментальных методов анализа: газовой и жидкостной хроматографии, атомной абсорбционной спектроскопии, атомно-эмиссионной спектроскопии, спектроскопии комбинационного рассеяния света, масс-спектрометрии и др. Предпочтение отдается методам анализа, предусматривающим минимальную предобработку образцов или ее полное отсутствие, экспрессный и неразрушающий характер, простоту применения, минимальное использование дополнительных реагентов. Эти особенности характерны для ряда спектроскопических методов, в том числе для спектроскопии ультрафиолетового, видимого, ближнего и среднего инфракрасного диапазонов.

В настоящее время спектроскопия с применением многопараметрических методов широко используется для классификации вин в соответствии с их географическим происхождением и калибровки физико-химических параметров вин [1–3], хотя и уступает традиционным методам с точки зрения погрешности. Ранее [4] мы осуществили моделирование физико-химических характеристик молдавских вин с помощью проекции на латентные структуры (*PLS – projection to latent structures*) [5] широкополосных спектров пропускания в диапазоне 220–2500 нм.

Целью работы являлось применение методов многопараметрического анализа широкополосных спектров пропускания для калибровки физико-химических показателей вин и повышение точности этой калибровки с помощью выбора спектральных переменных. Для повышения точности калибровки приме-

нены интервальные методы многопараметрического анализа данных. Эти методы являются компромиссными с точки зрения выбора спектральных переменных между классическим спектроскопическим и многопараметрическим подходами. Первый подразумевает использование небольшого количества спектральных отсчетов на отдельных длинах волн или в спектральных интервалах. Достоинством такого подхода является четкий физический смысл моделирования. Многопараметрический спектральный анализ выделяет из широкополосных и мультиколлинеарных спектров малоразмерный базис для построения калибровочной модели, физический смысл которой почти всегда неочевиден. Компромиссный подход [6], когда из широкополосных спектров частично удаляются спектральные переменные, приводящие к переопределенности многопараметрических моделей, уменьшает влияние избыточности спектральных данных на качество калибровки.

Измерения референтных значений физико-химических параметров вин и спектров пропускания

Выборка исследуемых вин, методы определения референтных значений их физико-химических параметров, методики измерения спектров пропускания и используемая спектральная аппаратура подробно описаны в [4].

С помощью анализа суммарной объясненной дисперсии зарегистрированных спектров с помощью метода главных компонент [7] было установлено, что среди рассмотренных методов предобработки спектров (центрирование спектров – приведение к нулю среднего значения на каждой длине волны, сглаживание центрированных спектров полиномами до 3 степени включительно, вычисление первой производной центрированных спектров с последующим сглаживанием, преобразование стандартного отклонения случайной величины с нормальным распределением) лучшим является сглаживание фильтром Савицкого–Голея полиномом 3-й степени по 9 точкам. Анализ графиков нагрузок и счетов в главные компоненты показал, что часть спектров ближнего инфракрасного диапазона с длинами волн более 1500 нм неинформативна и может быть изъята из рассмотрения без потери качества многопараметрических моделей.

В дальнейшем в спектрах были сохранены только 2241 отсчетов.

После предобработки экспериментальных данных построение калибровочной модели между зарегистрированными спектрами пропускания и референтными значениями физико-химических параметров вин проводилось с помощью интервальных методов *PLS* [8–10].

Применение интервальных методов проекции на латентные структуры к спектрам пропускания для определения физико-химических параметров вин

Простейший алгоритм *PLS* с выбором спектральных переменных [8] реализуется путем разделения диапазона измерений на заданное количество неперекрывающихся спектральных интервалов одинаковой ширины и сравнение качества моделирования по каждому отдельному интервалу. Целевой функцией для оценки качества моделирования служит среднеквадратичная ошибка калибровки по проверочной выборке (*RMSEC* – *Root Mean Square Error of Calibration*). Спектральные интервалы могут последовательно объединяться с целью достижения минимальной величины целевой функции, а могут по одному изыматься из полного диапазона измерений [9]. В некоторых модификациях алгоритма интервал может изменять размер и сдвигаться по спектру, например в интервальном *PLS* по движущемуся окну или интервальном *PLS* по комбинации движущихся окон (*scmwiPLS* – *searching combination moving window iPLS*) [10].

Среди перечисленных интервальных методов в настоящей работе будем использовать *scmwiPLS*. Именно с помощью этого метода точность калибровки температуры по спектрам флуоресценции $\text{Yb}^{3+}:\text{CaF}_2$ была повышена более чем в 2 раза по сравнению с *PLS* по широкополосным спектрам [11], что было лучшим показателем среди рассмотренных интервальных методов, в том числе использующих генетический алгоритм и нейронные сети [12].

В разработанной нами модификации *scmwiPLS* начальным этапом является выбор положения первого интервала минимальной ширины с помощью метода интервального *PLS*. Следующий интервал последовательно сдвигается в пределах всего спектрального

диапазона измерений и объединяется с первым по минимальной величине среднеквадратичной ошибки предсказания величины калибруемого параметра. Так как зависимость среднеквадратичной ошибки калибровки по проверочной выборке от количества спектральных переменных в модели не является монотонной, расширение спектрального интервала моделирования необходимо продолжать до достижения всего диапазона измерений. Минимальная ширина сдвигающегося спектрального интервала, выражаемая в числе спектральных переменных, на единицу превышает количество латентных структур, учитываемых при моделировании.

Поскольку на этапе планирования эксперимента был осуществлен подбор ограниченного и небольшого количества молдавских вин, невозможно воспользоваться самым надежным способом уменьшить ошибку предсказания величины параметра с помощью использования большой и разнообразной обучающей выборки [13]. В рассматриваемом случае из имеющегося количества исследуемых вин будет произведен отбор образцов в проверочную выборку, а по оставшимся будет осуществляться построение многопараметрической калибровочной модели. Проверочная выборка должна достаточно равномерно покрывать весь диапазон возможного изменения параметра. Два наиболее часто используемых алгоритма формирования проверочной выборки – это равномерное покрытие области изменения параметра и метод Кеннарда–Стоуна [14], при котором первый образец выбирается в середине диапазона изменения параметра, а последующие – наиболее удаленными от уже выбранных. Лучшие результаты по точности моделирования показывает метод формирования проверочной выборки [11], базирующийся на применении иерархического кластерного анализа. Кластерный анализ позволяет разделить имеющиеся вина на непересекающиеся кластеры либо по величине определяемого параметра, либо по величинам счетов в главные компоненты зарегистрированных спектров. В проверочную выборку выделяются образцы, ближайшему к центру каждого кластера. Количество кластеров зависит от диапазона изменения и распределения величины калибруемого параметра. Для каждого калибруемого параметра будет

подбираться оптимальное количество кластеров в пределах от 4 до 10 для красных вин и от 4 до 7 для белых, которому будет соответствовать оптимальное количество латентных структур и оптимальная ширина спектрального интервала.

Результаты многопараметрического моделирования

Проиллюстрируем разработанный метод многопараметрической калибровки физико-химических параметров вин на примере определения спиртуозности. На рисунке 1 для выборки красных вин представлены зависимости среднеквадратичных ошибок кросс-валидации и предсказания (левая ось) и остаточного отклонения предсказания для проверочной выборки *RPDC* (*Residual Predictive Deviation for Calibration*) [15] и для обучающей выборки *RPDCV* (*Residual Predictive Deviation for Cross-Validation*) (правая ось) от числа образцов в проверочной выборке. Остаточные отклонения предсказания определяются как отношения среднеквадратичного отклонения параметра в выборке к среднеквадратичной ошибке предсказания величины.

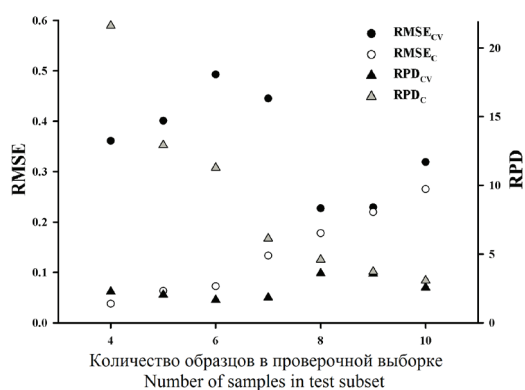


Рисунок 1 – Среднеквадратичные ошибки калибровки спиртуозности по проверочной выборке красных вин и кросс-валидации по обучающей выборке (левая ось) и остаточного отклонения предсказания

Figure 1 – Root mean square error of calibration of alcohol content in test subset of red wines and root mean square error of cross-validation in training subset (left axis) and residual predictive deviations

Видно, что 8 образцов красных вин из 21 формируют оптимальную проверочную выборку, основываясь на зависимости величины

RPDCV. Остальные 13 образцов формируют обучающую выборку. Моделирование по всему спектральному диапазону измерений (см. рисунок 2) показало, что минимальная величина остаточного отклонения предсказания для проверочной выборки соответствует *PLS* по 9 латентным структурам. Значит, в единичном интервале для применения *scmwiPLS* должно быть 10 спектральных переменных.

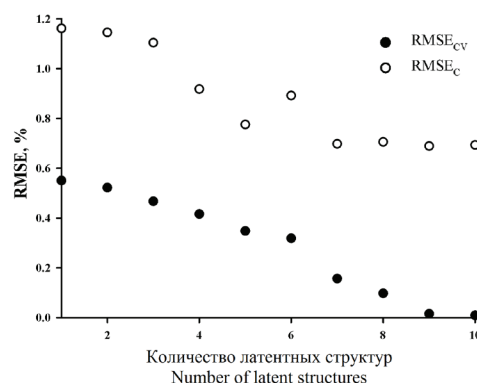


Рисунок 2 – Зависимости среднеквадратичных ошибок калибровки спиртуозности по проверочной выборке красных вин и кросс-валидации по обучающей выборке от количества латентных структур в широкополосной модели проекции на латентные структуры

Figure 2 – Root mean square error of calibration of alcohol content in test subset of red wines and root mean square error of cross-validation in training subset versus the number of latent structures in the wideband projection to latent structures model

Из 2241 спектральных переменных исключим 1 самую коротковолновую для того, чтобы оставшиеся 2240 можно было разделить на 224 неперекрывающихся интервала по 10 переменных. Далее применим метод *scmwiPLS* до моделирования по полному спектральному диапазону измерений. Для определения оптимальной комбинации спектральных переменных на рисунке 3 представлена зависимость *RMSEC* от количества используемых в модели интервалов.

Видно, что глобальный минимум *RMSEC* соответствует 54 интервалам или 540 спектральным переменным. График предсказанных величин спиртуозности красных вин в зависимости от соответствующих измеренных предсказан на рисунке 4.



Рисунок 3 – Зависимость среднеквадратичной ошибки калибровки величины спиртуозности красных вин от количества спектральных окон, используемых в модели интервальной проекции на латентные структуры по оптимальной комбинации движущихся окон

Figure 3 – Root mean square error of calibration of alcohol content in red wines versus the number of spectral windows used in the searching combination moving window interval projection to latent structures model

Величины среднеквадратичных ошибок калибровки спиртуозности по обучающей и проверочной выборкам при оптимальных параметрах моделирования методом *scmwiPLS* для красных вин равны 0,2 % об., а для белых – < 0,1 % об. и 0,1 % об., соответственно. При этом $RPDCV = 3,6$ и $RPDC = 4,6$ для

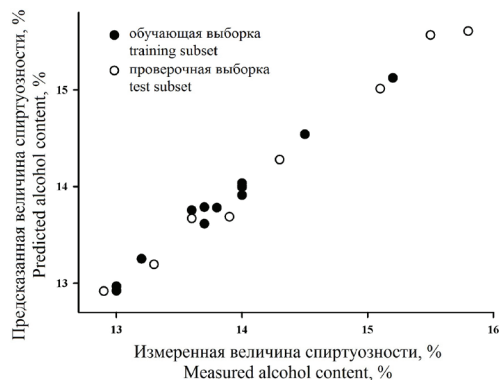


Рисунок 4 – Величины спиртуозности красных вин, предсказанные с помощью интервальной проекции на латентные структуры по комбинации движущихся окон, в зависимости от измеренных величин для обучающей и проверочной выборок при оптимальном выборе количества латентных структур и спектральных переменных

Figure 4 – Values of alcohol content in red wines predicted using searching combination moving window interval projection to latent structures versus the measured values for the training and test subsets with the optimal choice of the number of latent structures and spectral variables

красных вин и $RPDCV = 19,9$ и $RPDC = 7,1$ для белых. Эти показатели, превышающие пороговое значение 2,5 [15], характеризуют хорошее качество калибровки. В таблицах 1 и 2 представлены результаты моделирования химических характеристик красных и белых вин, для которых RPD также превышают значение 2,5.

Таблица 1/Table 1

Результаты калибровки характеристик красных вин, для которых остаточное отклонение предсказания превышает величину 2,5

Calibration results of the red wines characteristics, for which residual predictive deviation exceeds 2.5

	К	Са	Mg	Щавелевая кислота Oxalic acid	Яблочная кислота Malic acid	Янтарная кислота Succinic acid	2,3-бутиленгликоль 2,3-butylene glycol	Зола Ash	Фенольные соединения Phenolic compounds
$RMSE_{CV}$, мг/дм ³ mg/dm ³	29,5	2,0	0,6	38	26	109	17,3	100	155
$RMSE_C$, мг/дм ³ mg/dm ³	24,6	2,1	2,8	6	5	38	20,9	94	133
RPD_{CV}	2,6	3,7	13,8	4,7	2,9	2,7	3,2	3,6	2,9
RPD_C	3,1	3,6	3,0	2,8	15,6	7,7	2,6	3,8	3,3

Таблица 2/Table 2

Результаты калибровки характеристик белых вин, для которых остаточное отклонение предсказания превышает величину 2,5

Calibration results of the white wines characteristics, for which residual predictive deviation exceeds 2.5

	Mg	Винная кислота Tartaric acid	Лимонная кислота Citric acid	Молочная кислота Lactic acid	2,3-бутиленгликоль 2,3-butylene glycol	Зола Ash	Фенольные соединения Phenolic compounds	Растворимые соли Soluble salts
$RMSE_{CV}$, мг/дм ³ mg/dm ³	1,7	139	14	8	6,3	0,1	16,4	3,4
$RMSE_C$, мг/дм ³ mg/dm ³	0,9	165	23	3	5,4	0,1	2,0	31,6
RPD_{CV}	4,1	3,2	7,9	4,8	6,8	3,0	3,5	27,5
RPD_C	7,4	2,7	4,8	10,4	8,0	3,5	29,0	2,9

Удовлетворительное качество моделирования, при котором $2,0 < RPD < 2,5$ [18], достигнуто для калибровки содержания винной и яблочной кислот и глицерина для красных вин и винной кислоты для белых вин. Для калибровки содержания лимонной и молочной кислот, Na и растворимых солей для красных вин и щавелевой, яблочной и янтарной кислот, K, Na, Ca и глицерина в белых винах характерно неудовлетворительное качество многопараметрического моделирования с помощью *scmwiPLS*, при котором $RPD < 2,0$.

Заключение

Применение интервальных методов многопараметрического анализа данных позволяет существенно повысить точность калибровки физико-химических параметров вин по сравнению и с широкополосными многопараметрическими, и с классическими спектроскопическими методами. Например, среднеквадратичная ошибка калибровки спиртуозности красных вин методом интервальной проекции на латентные структуры по оптимальной комбинации движущихся окон меньше ошибки калибровки методом широкополосной проекции на латентные структуры в 6 раз и в 5 раз меньше среднеквадратичной ошибки калибровки по одиночному спектральному диапазону шириной 10 нм.

Таким образом, полученные результаты показывают, что спектроскопия ультрафиолетового, видимого и ближнего инфракрасного

диапазонов в сочетании с многопараметрическими методами обработки спектров является экспрессной, дешевой и неразрушающей технологией калибровки физико-химических параметров вин. Применение интервальных модификаций многопараметрических методов позволяет значительно увеличить точность калибровочных моделей.

Список использованных источников

1. *Dos Santos Claudia*. Merging vibrational spectroscopic data for wine classification according to the geographic origin / Claudia dos Santos [et al.] // Food Research International. – 2017. – Vol. 102. – P. 504–510. DOI: 10.1016/j.foodres.2017.09.018
2. *Ferreiro-Gonzalez, M.* FT-IR, Vis spectroscopy, color and multivariate analysis for the control of ageing processes in distinctive Spanish wines / M. Ferreiro-Gonzalez [et al.] // Food Chemistry. – 2019. – Vol. 277. – P. 6–11. DOI: 10.1016/j.foodchem.2018.10.087
3. *Dos Santos Claudia*. A review on the application of vibrational spectroscopy in the wine industry: From soil to bottle / C. dos Santos, R. Pascoa, J. Lopes // Trends Anal. Chem. – 2017. – Vol. 88. – P. 100–118. DOI: 10.1016/j.trac.2016.12.012
4. *Ходасевич, М.А.* Применение многопараметрического анализа широкополосных спектров пропускания для идентификации вин с географическим наименованием происхождения / М.А. Ходасевич [и др.] // Приборы и методы измерений. – 2016. – Т. 7, № 1. – С. 104–113. DOI: 10.21122/2220-9506-2016-7-1-104-113
5. *Pasquini, C.* Near infrared spectroscopy: A mature analytical technique with new perspectives –

A review / C. Pasquini // *Analytica Chimica Acta*. – 2018. – Vol. 1026. – P. 8–36. DOI: 10.1016/j.aca.2018.04.004

6. Mehmood, T. A review of variable selection methods in Partial Least Squares Regression / T. Mehmood [et al.] // *Chem. and Intel. Lab. Sys.* – 2012. – Vol. 118. – P. 62–69.

DOI: 10.1016/j.chemolab.2012.07.010

7. Esbensen, K.H. Principal Component Analysis: Concept, Geometrical Interpretation, Mathematical Background, Algorithms, History, Practice / K.H. Esbensen, P. Geladi // *Comprehensive Chemometrics* – 2009. – Vol. 2. – P. 211–226.

DOI: 10.1016/B978-044452701-1.00043-0

8. Norgaard, L. Interval Partial Least-Squares Regression (iPLS): A Comparative Chemometric Study with an Example from Near-Infrared Spectroscopy / L. Norgaard [et al.] // *Appl. Spectr.* – 2000. – Vol. 54. – P. 413–419. DOI: 10.1366/0003702001949500

9. Zou, X. Selection of the efficient wavelength regions in FT-NIR spectroscopy for determination of SSC of ‘Fuji’ apple based on BiPLS and FiPLS models / X. Zou, J. Zhao, Y. Li // *Vibr. Spectr.* – 2007. – Vol. 44. – P. 220–227. DOI: 10.1016/j.vibspec.2006.11.005

10. Du, Y.P. Spectral regions selection to improve prediction ability of PLS models by changeable size moving window PLS and searching combination moving window PLS / Y.P. Du [et al.] // *Anal. Chim. Acta.* – 2004. – Vol. 501, iss. 2. – P. 183–191.

DOI: 10.1016/j.aca.2003.09.041

11. Ходасевич, М.А. Выбор спектральных переменных и повышение точности калибровки температуры методом проекции на латентные структуры по спектрам флуоресценции $\text{Yb}^{3+}:\text{CaF}_2$ / М.А. Ходасевич, В.А. Асеев // *Оптика и спектроскопия*. – 2018. – Т. 124, № 5. – С. 713–717.

DOI: 10.1134/S0030400X18050089

12. Ходасевич, М.А. Калибровка температуры по спектрам флуоресценции допированного эрбием свинцово-фторидного стекла / М.А. Ходасевич [и др.] // *Оптика и спектроскопия*. – 2019. – Т. 126, № 3. – С. 295–298.

13. Anderson, R.B. Clustering and training set selection methods for improving the accuracy of quantitative laser induced breakdown spectroscopy / R.B. Anderson [et al.] // *Spectrochimica Acta Part B*. – 2012. – Vol. 70. – P. 24–32.

DOI: 10.1016/j.sab.2012.04.004

14. Nawar, S. Optimal sample selection for measurement of soil organic carbon using online vis-NIR spectroscopy / S. Nawar, A.M. Mouazen // *Computers and Electronics in Agriculture*. – 2018. – Vol. 151. – P. 469–477. DOI: 10.1016/j.compag.2018.06.042

15. Zornoza, R. Near infrared spectroscopy for determination of various physical, chemical

and biochemical properties in Mediterranean soils / R. Zornoza [et al.] // *Soil Biology & Biochemistry*. – 2008. – Vol. 40. – P. 1923–1930.

DOI: 10.1016/j.soilbio.2008.04.003

References

1. C. dos Santos, Pascoa Ricardo N.M.J., Sarraguca Mafalda Cruz, Porto Patricia A.L.S., Cerdeira Antonio L., Gonzalez-Saiz J.M., Pizarro C., Lopes Joao A. Merging vibrational spectroscopic data for wine classification according to the geographic origin. *Food Research International*, 2017, vol. 102, pp. 504–510.

DOI: 10.1016/j.foodres.2017.09.018

2. Ferreiro-Gonzalez M., Ruiz-Rodriguez Ana, Barbero Gerardo F., Ayuso Jesus, Alvarez Jose A., Palma Miguel, Barroso Carmelo G. FT-IR, Vis spectroscopy, color and multivariate analysis for the control of ageing processes in distinctive Spanish wines. *Food Chemistry*, 2019, vol. 277, pp. 6–11.

DOI: 10.1016/j.foodchem.2018.10.087

3. Dos Santos C. Pascoa R., Lopes J. A review on the application of vibrational spectroscopy in the wine industry: From soil to bottle. *Trends Anal. Chem.*, 2017, vol. 88, pp. 100–118. DOI: 10.1016/j.trac.2016.12.012

4. Khodasevich M., Scorbanov E., Obade L., Degtyar N., Cambur E., Rogovaya M. [Application of multivariate analysis of transmission spectra to identify wines with protected geographical indication (IGP)]. *Devices and Methods of Measurements*, 2016, vol. 7, no. 1, pp. 104–113 (in Russian).

DOI: 10.21122/2220-9506-2016-7-1-104-113

5. Pasquini C. Near infrared spectroscopy: A mature analytical technique with new perspectives – A review. *Analytica Chimica Acta*, 2018, vol. 1026, pp. 8–36.

DOI: 10.1016/j.aca.2018.04.004

6. Mehmood T., Hovde Liland Kristian, Snipen Lars, Saebo Solve. A review of variable selection methods in Partial Least Squares Regression. *Chem. and Intel. Lab. Sys.*, 2012, vol. 118, pp. 62–69.

DOI: 10.1016/j.chemolab.2012.07.010

7. Esbensen K.H., Geladi P. Principal Component Analysis: Concept, Geometrical Interpretation, Mathematical Background, Algorithms, History, Practice. *Comprehensive Chemometrics*, 2009, vol. 2, pp. 211–226.

DOI: 10.1016/B978-044452701-1.00043-0

8. Norgaard L., Saudland A., Wagner J., Nielsen J.P., Munck L., Engelsen S.B. Interval Partial Least-Squares Regression (iPLS): A Comparative Chemometric Study with an Example from Near-Infrared Spectroscopy. *Appl. Spectr.*, 2000, vol. 54, pp. 413–419.

DOI: 10.1366/0003702001949500

9. Zou X. Zhao J., Li Y. Selection of the efficient wavelength regions in FT-NIR spectroscopy for determination of SSC of ‘Fuji’ apple based on BiPLS and FiPLS

models. *Vibr. Spectr.*, 2007, vol. 44, pp. 220–227.

DOI: 10.1016/j.vibspec.2006.11.005

10. Du Y.P., Liang Y.Z., Jiang J.H., Berry R.J., Ozaki Y. Spectral regions selection to improve prediction ability of PLS models by changeable size moving window PLS and searching combination moving window PLS. *Anal. Chim. Acta*, 2004, vol. 501, pp. 183–191.

DOI: 10.1016/j.aca.2003.09.041

11. Khodasevich M.A., Aseev V.A. Selection of Spectral Variables and Improvement of the Accuracy of Calibration of Temperature by Projection onto Latent Structures Using the Fluorescence Spectra of Yb³⁺:CaF₂. *Optics and Spectroscopy*, 2018, vol. 124, pp. 748–752.

DOI: 10.1134/S0030400X18050089

12. Khodasevich M.A. Temperature calibration using fluorescence spectra of erbium-doped lead-fluoride glass. *Optics and Spectroscopy*, 2019, vol. 126, no. 3, pp. 216–219. **DOI:** 10.1134/S0030400X1903010X

13. Anderson R.B., James F. Bell III, Wiens Roger C., Morris Richard V., Clegg Samuel M. Clustering and training set selection methods for improving the accuracy of quantitative laser induced breakdown spectroscopy. *Spectrochimica Acta Part B.*, 2012, vol. 70, pp. 24–32.

DOI: 10.1016/j.sab.2012.04.004

14. Nawar S., Mouazen A.M. Optimal sample selection for measurement of soil organic carbon using online vis-NIR spectroscopy. *Computers and Electronics in Agriculture*, 2018, vol. 151, pp. 469–477.

DOI: 10.1016/j.compag.2018.06.042

15. Zornoza R., Guerrero C., Mataix-Solera J., Scow K.M., Arcenegui V., Mataix-Beneyto J. Near infrared spectroscopy for determination of various physical, chemical and biochemical properties in Mediterranean soils. *Soil Biology & Biochemistry*, 2008, vol. 40, pp. 1923–1930.

DOI: 10.1016/j.soilbio.2008.04.003