

## АНАЛИЗ ДОСТОВЕРНОСТИ КАЛИБРОВОЧНЫХ МОДЕЛЕЙ ДЛЯ ОПРЕДЕЛЕНИЯ СЫРОГО ПРОТЕИНА, СЫРОЙ КЛЕТЧАТКИ И СЫРОГО ЖИРА В СИЛОСЕ КУКУРУЗНОМ МЕТОДОМ БИК-СПЕКТРОМЕТРИИ (NIRS)

С. Н. МИХАЛЬЧУК, М. А. ПАСТУХОВА

Государственное научное учреждение «Полесский аграрно-экологический институт  
Национальной академии наук Беларуси»,  
г. Брест, Республика Беларусь, 224030, e-mail: pastukhova.marina@inbox.ru

(Поступила в редакцию 30.06.2023)

Представлены результаты валидации и параметры достоверности калибровочных моделей, разработанных Отраслевой научно-исследовательской лабораторией качества кормов Полесского аграрно-экологического института НАН Беларуси для определения сырого протеина, сырой клетчатки и сырого жира в кукурузном силосе методом NIRS. Калибровочные модели являются первой Брестской региональной библиотекой спектров силоса кукурузного, созданные с целью применения метода БИК-анализа (NIRS) в повседневной работе лаборатории в качестве альтернативного «мокрой химии» метода. В качестве референтного метода выбрана «мокрая химия». Анализ проводили в Отраслевой научно-исследовательской лаборатории качества кормов.

Установлены диапазоны спектров консервированных кормов для каждой из созданных моделей: для сырого протеина – 1187–2055 нм; для сырой клетчатки – 978–2480 нм; для сырого жира – 1387–2342 нм. Представлены комбинации параметров предварительной обработки спектров для каждой из описываемых калибровочных моделей. Приводится сравнительная характеристика значений показателей, полученных в результате анализа кормов двумя методами. В результате валидации калибровочных моделей установлен высокий коэффициент детерминации: для сырого протеина 0,89; для сырой клетчатки 0,98; для сырого жира 0,84. Разработанные калибровочные модели характеризуются низкой систематической ошибкой прогноза (bias для сырого протеина составляет (-1,4 г/кг СВ); для сырой клетчатки (-0,35 г/кг СВ); для сырого жира (-0,05 г/кг СВ). Стандартная ошибка предсказания модели для определения сырого протеина составляет 3,06; для сырой клетчатки 5,47; для сырого жира 2,14. Ошибка аппроксимации моделей колеблется от 2,4 % до 3,7 %. Таким образом подтверждается достоверность и возможность применения созданных калибровочных моделей для использования метода NIRS в ежедневной работе лаборатории.

**Ключевые слова:** БИК-анализ, калибровочная модель, спектр, диапазон, протеин, клетчатка, жир.

The results of validation and reliability parameters of calibration models developed by the Industrial Research Laboratory of Feed Quality of the Polesie Agrarian-Ecological Institute of the National Academy of Sciences of Belarus for the determination of crude protein, crude fiber and crude fat in corn silage using the NIRS method are presented. Calibration models are the first Brest regional library of corn silage spectra, created for the purpose of applying the NIR analysis method (NIRS) in the daily work of the laboratory as an alternative to “wet chemistry” method. “Wet chemistry” was chosen as the reference method. The analysis was carried out at the Industry Research Laboratory of Feed Quality.

Spectral ranges for conserved feed have been established for each of the created models: for crude protein – 1187–2055 nm; for crude fiber – 978–2480 nm; for crude fat – 1387–2342 nm. Combinations of spectra preprocessing parameters for each of the described calibration models are presented. A comparative description of the indicator values obtained as a result of feed analysis by two methods is provided. As a result of validation of calibration models, a high coefficient of determination was established: for crude protein 0.89; for crude fiber 0.98; for crude fat 0.84. The developed calibration models are characterized by a low systematic forecast error (bias for crude protein is 1.4 g/kg DM; for crude fiber 0.35 g/kg DM; for crude fat 0.05 g/kg DM). The standard prediction error of the model for the determination of crude protein is 3.06; for crude fiber 5.47; for crude fat 2.14. The approximation error of the models ranges from 2.4 % to 3.7 %. This confirms the reliability and the possibility of using the created calibration models to use the NIRS method in the daily work of the laboratory.

**Key words:** NIR analysis, calibration model, spectrum, range, protein, fiber, fat.

### Введение

Современное животноводство Брестской области находится на высоком технологическом уровне производства. По статистике, представленной комитетом сельского хозяйства и продовольствия Брестского облисполкома, в структуре товарной продукции сельского хозяйства животноводство занимает более 80 %; в том числе молочная продукция обеспечивает около 64 %, мясная – 32 %, другая побочная продукция – 4 % от общего объема производства.

Одним из определяющих факторов организации производства конкурентоспособной продукции животноводства является контроль качества кормов. В структуре себестоимости готовой продукции корма занимают 60–65 % [1]. Поэтому контроль качества кормов с учетом фактического содержания, переваримости и усвояемости элементов питания является одним из важнейших факторов снижения издержек, в числе которых рациональное использование кормов, с одной стороны, и повышение эффективности технологии содержания животных при сохранении их здоровья и увеличения продолжительности хозяйственного использования, с другой [2, 3].

В Республике Беларусь ежегодно анализируются десятки тысяч образцов кормов. На сегодняшний день регламентирующими документами по проведению соответствующих анализов являются ГОСТы; методами исследования – традиционная «мокрая химия», основанная на химическом взаимодействии веществ. Наряду с высокой достоверностью получаемых результатов «мокрая химия» обладает рядом недостатков: длительное время анализа, высокая себестоимость, использование химических веществ и проблема их утилизации с отходами. В этой связи исследования, представленные в статье, являются актуальными с точки зрения применения в Республике Беларусь более экономичного и экологичного метода исследования – БИК-анализа (NIRS).

БИК-анализ (NIRS), или спектроскопия ближнего инфракрасного отражения (Near-Infrared Reflectance Spectroscopy) – это метод физического анализа, который использует ближнюю область инфракрасного излучения. Это современный инструментальный метод количественного и качественного анализа различных объектов, основанный на сочетании спектроскопии и статистических методов исследования многофакторных зависимостей. Метод основан на том, что спектры поглощения молекул являются характеристическими величинами для данного вещества, а интенсивность поглощения связана с содержанием поглощающего компонента в облучаемом объекте [4].

Методы, основанные на спектральном анализе в ближней инфракрасной области (БИК (NIRS)-анализ), широко применяются во многих странах мира для оперативного анализа целого ряда показателей качества сельскохозяйственной продукции. Опыт исследований кормов методом NIRS в мировой практике насчитывает более чем 90-летнюю историю. Одной из первых лабораторий, использующих NIRS для анализа кормов, является лаборатория компании Eurofins Agro [5]. Одно из основных преимуществ данного метода – минимизация риска влияния человеческого фактора на результат анализа, а также минимальная пробоподготовка, ограничивающаяся сушкой и измельчением анализируемого материала.

Использование метода БИК-анализа возможно только при наличии обширной базы данных референтных образцов (библиотеки спектров и фактических значений показателей), проанализированных химическими методами. На основе этих баз данных создаются калибровочные модели, позволяющие интерпретировать полученные методами БИК-спектроскопии данные. Опыт сотрудничества с голландской фирмой Eurofins Agro и доступ к глобальной библиотеке спектров, позволил ОНИЛКК самостоятельно построить первые в Республике Беларусь адаптированные под региональные условия калибровочные модели для различных групп кормов с учетом их ботанического состава. Использование двух подходов к анализу кормов («мокрая химия» и NIRS-анализ) позволяет помимо сокращения времени анализа осуществлять периодическую верификацию (контрольную проверку) полученных данных, что является гарантией достоверности исследований ОНИЛКК.

#### **Основная часть**

Настоящие исследования проводились с целью подтверждения достоверности результатов определения сырого протеина, сырого жира и сырой клетчатки в силосе кукурузном, полученных методом БИК-спектрометрии при помощи созданных калибровочных моделей в сравнении с результатами, полученными методами «мокрой химии».

Для выполнения БИК-анализа (NIRS) требуется предварительная градуировка (калибровка) по данным референтного химического анализа и спектрам этих же референтных образцов. Важнейшее значение при БИК-анализе имеет соблюдение единообразия в процессе пробоподготовки. Сушка образцов проводилась при температуре 70 °С до постоянной массы; измельчение – на мельнице Fritsch Pulverisette 19 с использованием сита с размером ячейки 0,5 мм.

Спектральные данные образцов получены с помощью прибора Фурье спектрометр инфракрасный Quant (Q-interline, Дания).

Референтным методом для определения сырого протеина являлся метод Кьельдаля (ГОСТ 13496.4–2019); анализ проводился на приборе Gerhardt Vapodest 45s; сырой клетчатки – ГОСТ 13496.2–91 на приборе Gerhardt Fibretherm; сырой жир экстрагировали методом Сокслета по ГОСТ 13496.15–2016.

Создание калибровочных моделей осуществлялось в программе Grams IQ версии 9.3 от Thermo Fisher Scientific Inc., работу со спектрами проводили в программе Spekwin32 версии 1.72.0. Спектры образцов получали с помощью программы InfraQuant версии 4.0.20013.2 от фирмы q-Interline (Дания).

Для разработки калибровок использовался метод наименьших квадратных (PLS-1). Предварительная обработка спектральных данных образцов и их точных данных заключалась в выборе диапазона спектра с наиболее интенсивными полосами поглощения, характерными для данного пока-

зателя; выборе типа диагностики; производной алгоритма Савицкого-Голая (SG) для устранения шума и сглаживания кривых в данных [4], а также количества точек сглаживания длин волн.

При построении модели для каждого из трех показателей (сырой протеин, сырой жир, сырая клетчатка) использовались различные комбинации параметров предварительной обработки спектров. Из них были выбраны параметры, обеспечивающие построение калибровочной модели с высокой корреляционной зависимостью предсказанных и истинных значений (табл. 1). В результате работы со спектрами образцов в генеральной совокупности (371 образец) определены «выбросы»: по сырому протеину 54 образца, по сырой клетчатке 33 образца, по сырому жиру – 45. Для валидации использовали 20 образцов, которые не участвовали в создании моделей.

Таблица 1. Параметры предварительной обработки спектральных данных

Параметры	Протеин	Клетчатка	Жир
Метод калибровки	PLS-1	PLS-1	PLS-1
Тип диагностики	перекрестная проверка	перекрестная проверка	перекрестная проверка
Тип производной	SG 1st	SG 2nd	SG 1st
Количество точек производной	5	5	5
Коррекция длины пути	MSC correction	–	–
Спектральный диапазон, нм	1357–2618	1310–2656	1135–2608
Коэффициент детерминации (R <sup>2</sup> )	0,95	0,93	0,91

Для калибровочной модели по определению сырого протеина был применен метод устранения эффектов рассеяния – MSC (Multiplicative Scatter Correction). В данном случае это было вызвано необходимостью корректировки спектра устранением эффектов рассеяния света и других источников шума, регистрирующихся прибором [6].

Для диагностики созданных моделей использовался тип валидации – перекрестная проверка (Cross Validation). Перекрестная проверка – это способ формирования прогнозных статистических данных, где на постоянной основе множество образцов удаляется из калибровочной совокупности; модель рассчитывается на основе остальных образцов [6]. Последовательное удаление образца или группы образцов обусловлено отличием спектров этих образцов от остального множества калибровочной выборки. В свою очередь эти отличия могут возникать вследствие неправильного отбора пробы, плохой сушки образца, размола, и температуры образца, несоответствующей температуре среды, в которой выполняется анализ. Для определения подобных образцов мы использовали критерий расстояния Махаланобиса.

Считается, что NIR-область охватывает диапазон света с длинами волн от 750 до 2500 нм [7]. Спектры в этой области состоят из обертонов и комбинированных полос. Некоторые приборы обладают возможностью измерений только в области обертонов. Используемый в наших исследованиях спектрометр Quant предоставляет информацию по всему спектру, что обеспечивает получение более точных и надежных результатов [7]. Поэтому при построении модели мы выбираем не только узкую область спектра, с наиболее интенсивной полосой поглощения, скажем для протеина, но всю информативную область – «область отпечатков пальцев» (рис. а, б).

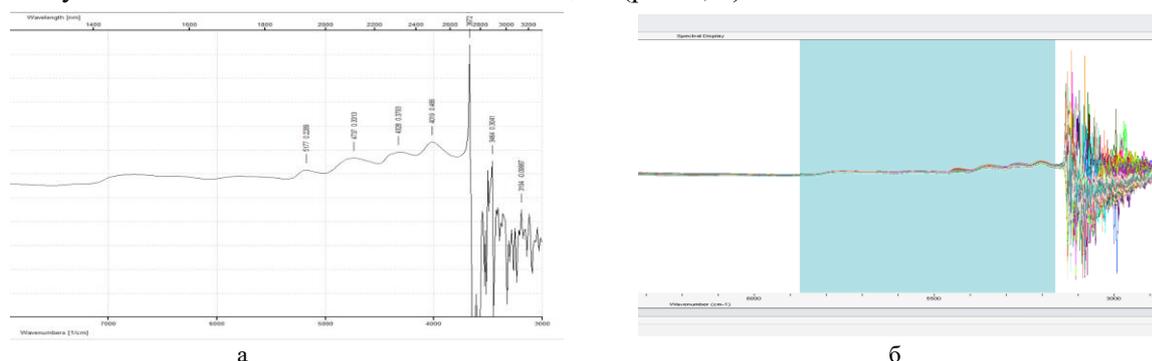


Рис. Область «отпечатков пальцев» спектров силоса кукурузного  
а – в программе Spekwin32; б – в программе Grams IQ

Повторяющийся характер спектров NIR всегда обеспечивает некоторую степень избыточности из-за того, что связи N-H, O-H и C-H поглощаются на разных длинах волн одновременно (табл. 2) [7]. Следствием этой избыточности является то, что мы можем извлечь одинаковую информацию на разных участках спектра. Это создает трудности в интерпретации БИК-спектров. Однако информация, спрятанная в вершинах и впадинах (на пиках) спектров, имеет специфический характер для каждого

типа органического материала. Если образец сенажа просканировать на модели, сделанной для силоса, то модель определит полученный спектр образца как «выброс» (outlier).

Таблица 2. Положение наиболее интенсивных полос поглощения

Сырой протеин		Сырая клетчатка		Сырой жир	
нм	см <sup>-1</sup>	нм	см <sup>-1</sup>	нм	см <sup>-1</sup>
1187	8424	978	10225	1387	7210
1485	6734	2268	4409	1722	5807
1690	5917	2480	4032	2306	4336
1972	5071			2342	4270
2055	4866				

Наиболее существенным фактором, влияющим на достоверность рисунка спектра, является пробоподготовка. Для БИК-анализа важно методично и единообразно готовить пробу образца, как для построения модели, так и для валидации. Лаборатория ОНИЛКК заранее определилась со стратегией пробоподготовки, принимая во внимание возможности мельницы для размола образцов (сито 0,5 мм) и специализированных сушильных шкафов ( $t=70^\circ$ ). Если образец силоса, размолотый на сите 1 мм, просканировать на модели для силоса, размолотого на сите 0,5 мм, то испытуемый образец, чаще всего, выйдет за установленный предел Махаланобиса и будет считаться «выбросом».

Для валидации и оценки прогностической способности калибровочных моделей мы использовали 20 независимых образцов кукурузного силоса (не входящих в состав калибровочного набора) с известными значениями показателей, полученными методами «мокрой химии». Образцы приготовлены и проанализированы в соответствии с пробоподготовкой калибровочной выборки (табл. 3).

Таблица 3. Результаты валидации полученных моделей

Протеин, г/кг СВ (55–120)*			Клетчатка, г/кг СВ (140–300)*			Жир, г/кг СВ (20–50)*		
NIRS	Мокрая химия	Разница	NIRS	Мокрая химия	Разница	NIRS	Мокрая химия	Разница
71	75	-4	193	193	0	35	36	-1
77	81	-4	168	177	-9	41	39	2
99	105	-6	236	231	5	30	29	1
83	82	1	208	202	6	34	32	2
85	88	-3	223	216	7	36	34	2
81	77	4	202	198	4	39	36	3
80	78	2	168	169	-1	35	35	0
77	79	-2	241	233	8	35	33	2
95	95	0	195	200	-5	39	39	0
88	88	0	213	216	-3	36	38	-2
92	91	1	222	217	5	36	36	0
64	66	-2	151	158	-7	44	40	4
77	76	1	152	158	-6	38	38	0
87	85	2	231	230	1	32	32	0
84	86	-2	208	210	-2	37	40	-3
82	83	-1	209	204	5	33	35	-2
87	91	-4	252	259	-7	28	32	-4
89	91	-2	213	220	-7	31	32	-1
80	85	-5	270	267	3	24	27	-3
82	86	-4	265	269	-4	24	25	-1
Смещение ( <b>bias</b> ), г/кг СВ	-1,40	Смещение ( <b>bias</b> ), г/кг СВ	-0,35	Смещение ( <b>bias</b> ), г/кг СВ	-0,05			
Станд. отклонение ( <b>SD</b> ), г/кг СВ	2,70	Станд. отклонение ( <b>SD</b> ), г/кг СВ	5,46	Станд. отклонение ( <b>SD</b> ), г/кг СВ	2,14			
Станд. ошибка предсказания ( <b>SEP</b> ), г/кг СВ	3,06	Станд. ошибка предсказания ( <b>SEP</b> ), г/кг СВ	5,47	Станд. ошибка предсказания ( <b>SEP</b> ), г/кг СВ	2,14			
Коэффициент детерминации ( <b>R<sup>2</sup></b> )	0,89	Коэффициент детерминации ( <b>R<sup>2</sup></b> )	0,98	Коэффициент детерминации ( <b>R<sup>2</sup></b> )	0,84			
Ошибка аппроксимации, %	3,1	Ошибка аппроксимации, %	2,4	Ошибка аппроксимации, %	3,7			

\* – диапазоны концентраций (г/кг), ограничивающие калибровочную модель (оптимальные значения показателя).

Оценка ошибок имеет первостепенное значение для обеспечения качества БИК-анализа. Прогностическая способность и достоверность калибровочной модели определялась нами следующими статистическими показателями: стандартное отклонение (**SD**), коэффициент детерминации (**R<sup>2</sup>**), стандартная ошибка предсказания (**SEP**, ошибка аппроксимации, систематическая ошибка прогноза (**bias**).

**SEP** представляет собой среднеквадратическое отклонение от данных химического анализа и является мерой разброса между истинными и прогнозными значениями, полученными от модели. Эта

величина характеризует суммарную погрешность метода и включает как случайную, так и систематическую ошибку метода [4]. Чем ниже значение SEP, тем точнее модель прогнозирует содержимое неизвестного образца [6, 7]. Показатель является ориентиром при обновлении модели, когда мы хотим улучшить ее прогностическую способность. Каждая последующая дополненная новыми эталонными образцами модель должна иметь SEP ниже предыдущей. Так мы понимаем, что модель обучается прогнозировать статистически достоверные данные.

Ошибка аппроксимации выражается в абсолютных величинах, либо в процентах. Определяет вероятную ошибку выявленного уравнения регрессии, основанного на оценке фактических значений каждого исследуемого образца. Рабочей и достоверной может считаться модель, у которой ошибка аппроксимации  $<10$ .

Важное значение для оценки достоверности модели имеет отношение стандартного отклонения (SD) к ошибке предсказания (SEP) [10]. В идеальном варианте они должны быть равны или близки друг к другу, при высокой корреляционной зависимости двух методов анализа [10]. В представленных исследованиях наилучший результат получен в модели по определению сырой клетчатки: при почти равном SEP и SD коэффициент детерминации имеет значение, близкое к единице – 0,98. В моделях по определению сырого протеина и сырого жира отношение SEP и SD является близким и равным соответственно, однако коэффициенты детерминации меньше 0,90, хотя и находятся в пределах общепринятых допущений ( $>0,70$ ).

Значение bias (смещение) интерпретируется как систематическая ошибка прогноза [8, 11]. Bias рассчитывается как средняя разница между методами БИК-анализа и «мокрой химии». Значение bias равно -1,40 говорит о том, что модель по сырому протеину в среднем на 1,40 г/кг СВ занижает значения испытуемых образцов. Программа InfraQuant имеет возможность корректировать смещение. Мы пользуемся этой возможностью, только если по результатам валидации из всей выборки 20-ти образцов больше 80 % показывают положительную или отрицательную разности. В нашем случае с моделью по сырому протеину процентное соотношение положительных и отрицательных разностей равно 40:60 соответственно; поэтому, учитывая незначительное смещение (1,4 г/кг СВ) мы не вносим корректировку на смещение, а ставим в программе InfraQuant в графе bias по умолчанию «0», т.к. bias в идеальном варианте должен быть равен нулю. В двух других моделях мы, аналогично, не ставим поправку на смещение.

Интересным является метод валидации моделей у наших коллег из Дании фирмы q-Interline. Они выделяют термин «бюджет ошибок», который представляет собой величину SEP, учитывающую работу прибора NIR, точность референтного метода, а также влияние пробоподготовки [11]. Таким образом они оценивают вклад каждого фактора в общую ошибку метода.

Существуют и другие статистические показатели и методы оценки работы калибровочных моделей. Среди них выделяют SEC, SECV, RPD, SEP(C), Slope и другие [6, 7, 10, 12].

### **Заключение**

В соответствии с поставленной целью нами сделаны следующие выводы. Целесообразно использование БИК-анализа (NIRS) для оценки показателей качества травяных кормов как альтернативного «мокрой химии» метода исследований, являющегося экономичным и экологичным.

Определен спектральный диапазон с наиболее интенсивными полосами поглощения в области «отпечатков пальцев», которые целесообразно учитывать при построении калибровочных моделей: для сырого протеина – 1187–2055 нм; для сырой клетчатки – 978–2480 нм; для сырого жира – 1387–2342 нм.

В результате валидации калибровочных моделей установлен высокий коэффициент детерминации: для сырого протеина 0,89; для сырой клетчатки 0,98; для сырого жира 0,84. Модели характеризуются низкой систематической ошибкой прогноза (bias для сырого протеина составляет (-1,4 г/кг СВ); для сырой клетчатки (-0,35 г/кг СВ); для сырого жира – (0,05 г/кг СВ). Стандартная ошибка предсказания модели для определения сырого протеина составляет 3,06; для сырой клетчатки 5,47; для сырого жира 2,14. Ошибка аппроксимации моделей колеблется от 2,4 % до 3,7 %.

Вышеперечисленные показатели оценки достоверности созданных моделей свидетельствуют об их высокой надежности и пригодности для широкого практического использования.

### *ЛИТЕРАТУРА*

1. Шупик, М. В. Кормление сельскохозяйственных животных / М. В. Шупик, А. Я. Райхман. – Горки: БГСХА, 2014 – 236 с.
2. Рядчиков, В. Г. Основы питания и кормления сельскохозяйственных животных: учеб.-практ. пособие / В. Г. Рядчиков. – Краснодар: КубГАУ, 2012. – 328 с.

3. Пастухова, М. А. Сравнительная оценка силоса из силфии пронзеннолистной, кукурузы и сорго / М. А. Пастухова, Б. В. Шелюто // Технологические аспекты возделывания сельскохозяйственных культур: сб. ст. по материалам XI научно-практической конф., г. Горки, 29–30 января 2018 г. / Белорусская гос. сельскохозяйственной академия; редкол.: А. С. Мастеров [и др.] – Горки, 2018. – С. 186–190.
4. Крищенко, В. П. Ближняя инфракрасная спектроскопия: научное издание / В. П. Крищенко. – Москва: научно-методический центр по инфракрасной спектроскопии АО «Интерагротех», 1997. – 639 с.
5. Еврофинс-Агро [Электронный ресурс]. – Режим доступа: <https://www.eurofins.com/>. – Дата доступа: 19.05.2023.
6. Обеспечение качества работы лабораторий по анализу кормов для животных / Джим Белтроп [и др.]. – Руководство ФАО по вопросам животноводства и здоровья животных, № 14. – Рим, 2013. – 226 с.
7. A guide to near infrared technology [Electronic resource]. – Mode of access: <https://www.abvista.com/news/a-guide-to-nir-understanding-nir-spectra>. – Date of access: 26.06.2023.
8. Direct and indirect means of predicting forage quality through near infrared reflectance spectroscopy [Electronic resource]: Field crops research. – 2003. – Mode of access: <https://www.sciencedirect.com/science/article/abs/pii/S0378429003001400>. – Date of access: 26.06.2023.
9. Knadel, M. VIS/NIR mapping of TOC and extent of organic soils in the Nørre Å valley / M. Knadel, M.H. Greve, A. Thomsen // NJF Seminar 413: agricultural applications of NIRS and NIT, Aarhus University. – Slagelse, Denmark, 2009. – P. 10.
10. Shenk, J. S. Application of NIR spectroscopy to agricultural products. Handbook of near-infrared analysis / J. S. Shenk. – Boca Raton, Florida, U.S.A. – 2008. – 8 с.
11. Calibration and validation guideline [Electronic resource]: Technical note, 2017. – Mode of access: <https://q-interline.com/technology/statistics>. – Date of access: 26.06.2023.
12. Prediction of fatty acid and mineral composition of lentils using near infrared spectroscopy [Electronic resource]: Journal of food composition and analysis. – 2021. – Mode of access: <https://www.sciencedirect.com/science/article/pii/S0889157521002234/>. – Date of access: 26.06.2023.