

Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Российский экономический университет имени Г.В. Плеханова»

На правах рукописи

Метленкин Дмитрий Андреевич

**Разработка методических подходов применения оптической спектроскопии и
гиперспектрального изображения для идентификации и контроля качества пи-
щевых продуктов**

4.3.3. Пищевые системы

диссертация на соискание ученой степени
кандидата технических наук

Научный руководитель –
доктор технических наук, доцент
Платов Ю.Т.

Москва – 2024

Оглавление

Введение.....	4
Глава 1 Мировые тенденции применения методов оптической спектроскопии и анализа изображений для идентификации и контроля качества пищевых продуктов	12
1.1 Теоретические аспекты методов оптической спектроскопии и методов анализа изображений	12
1.2 Хемометрические методы анализа спектральных данных	14
1.3 Применение методов оптической спектроскопии и методов анализа изображений для идентификации и контроля качества пищевых продуктов	21
1.3.1 Продукты зерновых культур.....	21
1.3.2 Животные и растительные масла	27
1.3.3 Плодоовощная продукция.....	31
Глава 2 Объекты и методы исследования	36
2.1 Постановка научной задачи	36
2.2 Характеристика объектов исследования	41
2.3 Методы проведения исследований	44
2.3.1 Инструментальные методы	45
2.3.2 Методы многомерного анализа	52
Глава 3 Разработка моделей классификации пищевых продуктов на категории качества	55
3.1 Градация гречневой крупы по данным спектроколориметрии методом дискриминантного анализа	55
3.2 Градация гречневой крупы по <i>UV-VIS-NIR</i> -спектрам и цветовым координатам.....	61

3.3	Идентификация зернового кофе методами ИК-Фурье спектроскопии и многомерного анализа	74
	Глава 4 Разработка калибровочных моделей определения показателей качества пищевых продуктов сочетанием методов ИК-спектроскопии и многомерного анализа	83
4.1	Идентификация и подтверждение подлинности (аутентичности) сливочного масла при возможной фальсификации пальмовым маслом.....	83
4.2	Прогнозирование кислотного числа жира и сроков хранения гречневой крупы.....	89
	Глава 5 Контроль качества пищевых продуктов методами анализа гиперспектральных изображений.....	92
5.1	Сортировка зерна гречихи с помощью гиперспектральных изображений....	92
5.2	Классификация плодов авокадо Хасс по наличию дефектов с помощью гиперспектральных изображений.....	97
5.3	Сортировка плодов авокадо по гиперспектральным изображениям.....	111
5.4	Применение вегетационных индексов гиперспектральных изображений для сортировки плодов авокадо.....	125
	Заключение	135
	Список сокращений и условных обозначений.....	137
	Список использованной литературы.....	139
	Приложение А (справочное).....	164
	Приложение Б (обязательное).....	165
	Приложение В (обязательное)	166

Введение

Актуальность темы. Подтверждение подлинности и контроль качества пищевых товаров осуществляются на основе органолептических и инструментальных методов анализа, включая ядерный магнитный резонанс (ЯМР), масс-спектрометрия (МС), газовая и жидкостная хроматография (ГХ и ЖХ), полимеразная цепная реакция (ПЦР) и других (G.Artavia et al., 2021). Указанные методы являются высокоточными и востребованными, но трудоемкие, разрушающие, а для их применения необходима высокая квалификация оператора.

Поэтому большое практическое применение имеют неразрушающие методы контроля, учитывая их метрологические и эксплуатационные характеристики (информативность, оперативность, чувствительность, точность и стоимость анализа), которые позволяют оценить качество быстро, в режиме реального времени и на месте (E.Mendes, N.Duarte, 2021).

Среди неразрушающих методов наиболее востребованы методы оптической спектроскопии и анализа изображений. Для мониторинга качества пищевых продуктов распространено использование спектроскопии в инфракрасной и ближней инфракрасной области (ИК-, БИК-спектроскопия), спектроскопии в видимом диапазоне, включая спектрофотометрию и спектроколориметрию, а также спектроскопию комбинационного рассеяния. К перспективным направлениям относят машинное зрение, мультиспектральные и гиперспектральные изображения (*HSI*) в видимом и ближнем инфракрасном диапазоне, полученные с помощью цифровых камер (H.O.Velesaca et al., 2021). Гиперспектральное изображение обладает преимуществами оптической спектроскопии и предоставляет спектрально-пространственную информацию об объектах исследования (M.Manley, 2014). При этом преимуществом таких методов является их портативность, что увеличивает число возможных приложений (дистанционная съемка, установка камер на сортировочные линии).

В результате сбора информации формируется база данных: *2D*- или *3D*-матрицы, на основании которых с помощью многомерных методов строятся компьютерные модели (О.Е.Родионова и А.Л.Померанцев, 2006; Ю.Б.Монахова, С.П.Муштакова, 2017). Такие компьютерные модели позволяют решать две основные задачи: исследование данных с целью классификации и дискриминации объектов; прогнозирование новых значений (О.Е.Родионова и А.Л.Померанцев, 2006).

Сочетание методов оптической спектроскопии, гиперспектрального изображения совместно с многомерными методами является перспективным направлением для оперативного контроля пищевых продуктов на всех этапах технологического жизненного цикла товаров (ISO 14001).

Степень разработанности темы исследования. Исследования по разработке и применению методов оптической спектроскопии и гиперспектральных изображений совместно с методами многомерного анализа проведены многими отечественными и зарубежными учеными, а именно: О.Е.Родионовой, А.Л.Померанцевым, Ю.Б.Монаховой, Ю.В.Зонтовым, J.Straus, S.Kreft, K.M.G.Lima, V.Baeten, L.E.Rodriguez-Saona и другими. Известны работы по анализу гиперспектральных изображений для идентификации и классификации продуктов питания, которые проводились Б.М.Шурыгиным, А.Е.Соловченко, E.Hurtado-Fernández, J.Pinto, A.Anne Frank Joe, A.Gopal, H.Croft, B.Jia, W.Wang и другими.

Вместе с тем применение методов оптической спектроскопии и анализа изображений, включая неразрушающие методы и использование портативных камер, для мониторинга качества пищевых товаров изучено недостаточно. Для возможности цифровизации процедуры идентификации и контроля качества пищевых продуктов необходимо разработать методические подходы разработки компьютерных моделей с использованием методов оптической спектроскопии, анализа изображений и многомерного анализа.

Цели и задачи исследования. Цель диссертационной работы – разработка методических подходов применения оптической спектроскопии и гиперспектрального

изображения для идентификации и контроля качества отдельных видов пищевых продуктов.

Для осуществления цели поставлены и решены следующие задачи:

- научное и теоретическое обоснование методических подходов применения методов оптической спектроскопии, гиперспектральных изображений и многомерного анализа для разработки компьютерных моделей идентификации и контроля качества пищевой продукции;

- формирование и анализ базы спектральных данных: выявление характеристических частот спектра методом декомпозиции (метод главных компонент), связанных с физико-химическим состоянием (степень созревания, сроки хранения, наличие дефектов) отдельных видов пищевых продуктов;

- разработка и апробация моделей классификации пищевых товаров на категории качества: зернового кофе по ботаническому виду, обжарке и географическому региону произрастания; гречневой крупы по времени урожая (весенний, осенний) и в зависимости от гидротермической обработки;

- разработка компьютерных моделей мониторинга качества пищевых товаров (плоды авокадо по наличию дефектов и зрелости, зерно гречихи по степени выполненности) с применением гиперспектральных изображений в видимом и ближнем инфракрасном диапазоне;

- разработка и апробация калибровочных моделей прогнозирования свойств пищевых товаров (сливочное масло, гречневая крупа).

Научная новизна. Диссертационная работа содержит элементы научной новизны по пунктам 19, 29 паспорта специальности 4.3.3, а именно:

1) Разработан методический подход анализа гиперспектральных изображений, полученных использованием портативной (400-1000 нм) и лабораторной (935-1720 нм) камер, для контроля качества пищевой продукции:

- для обнаружения визуально не выявляемых дефектов авокадо. Впервые осуществлена визуализация гиперспектральных изображений изменения состояния плодов авокадо в процессе хранения с применением в качестве псевдоцветов длин волн спектра и главных компонент, определенных многомерными методами, в программном комплексе Альбеда 4.0.23. (МФТИ, Россия). Разработаны компьютерные модели классификации плодов авокадо по наличию дефектов методами *PLS-DA*, *SIMCA*;

- для определения влажности плодов авокадо посредством отбора спектральных сигнатур *HSI* случайным образом со всей поверхности плодов или всей поверхности изображения *HSI* плодов как области интереса (*ROI*) и компьютерного моделирования;

- для сортировки плодов авокадо по степени зрелости путем расчета и модификации вегетационных индексов (*PWI* и *NDVI757/679* соответственно) и визуализации гиперспектрального изображения по их значениям;

- для сортировки зерен гречихи по степени выполненности на основании отбора спектральных сигнатур из гиперспектрального изображения в диапазоне 935-1720 нм разработана классификационная модель методом *PLS-DA*.

2) Выявлены критерии идентификации – функциональные группы полос поглощения ИК-спектра: в диапазоне 1700-700 см⁻¹ для отличия зернового кофе от декофеинизированного кофе; при 2928, 2851, 1742, 1650, 1161 см⁻¹ (колебания С-Н, С=О, С-С связей), объясняющие отличия по составу сливочных, растительно-сливочных масел и сливочных масел с возможной фальсификацией пальмовым;

3) Выявлены критерии идентификации по данным *UV-VIS-NIR*-спектров и координат спектроколориметрической системы *CIE L*a*b** методами многомерного анализа гречневой крупы от времени сбора крупы (осень/весна), гидротермической обработки (ГТО) и срока хранения;

4) Разработаны компьютерные модели по ИК-спектрам: определения содержания пальмового масла в составе сливочного и растительно-сливочного масла методом

ПЛС; определения кислотного числа жира и сроков хранения гречневой крупы методом ПЛС.

5) Установлены зависимости изменения спектральных характеристик в диапазоне «красный край» (690-780 нм, *red edge*) и в диапазоне первого обертона колебаний ОН (900-970 нм) от степени созревания и изменения качества пищевой продукции (на примере плодов авокадо и гречневой крупы).

Теоретическая значимость. Разработаны и апробированы методические подходы применения оптической спектроскопии и гиперспектрального изображения для идентификации и контроля качества пищевых продуктов.

Выявлены полосы поглощения с максимальной дисперсией в видимой и инфракрасной областях спектра, являющиеся критериями идентификации исследованных пищевых продуктов.

Продемонстрирован подход обработки 3D-матриц данных на примере гиперспектральных изображений, включая сегментацию изображений, отбор спектральных данных, визуализацию с использованием псевдоцветов для разработки компьютерных моделей идентификации пищевых продуктов.

Практическая значимость. Разработаны методические подходы применения методов оптической спектроскопии и гиперспектрального изображения для идентификации и контроля качества пищевых продуктов на примере зернового кофе, гречневой крупы, сочной плодоовощной продукции и молочной жиросодержащей продукции.

Результаты работы прошли апробацию в лаборатории качества *Q-lab* сети «Азбука Вкуса», совместно с сотрудниками лаборатории были разработаны и применены хемометрические модели для определения степени зрелости и внутренних дефектов авокадо.

Результаты работы, связанные с разработкой и применением компьютерных моделей мониторинга качества пищевых товаров, используются при проведении лекци-

онных и практических занятий, подготовке курсовых и выпускных квалификационных работ бакалавров и магистров в рамках учебного процесса по направлениям подготовки «Товарный менеджмент и экспертиза», «Пищевые инновации и продакт-дизайн» ФГБОУ ВО «РЭУ им. Г.В. Плеханова».

Методология и методы исследования. Исследования проводились с использованием специально разработанных методических подходов, в которых предложено научное обоснование и разработка компьютерных моделей идентификации и контроля качества пищевых продуктов с применением инструментальных методов (оптическая спектроскопия, гиперспектральное изображение) и компьютерного моделирования, включающего построение баз данных, создание наборов данных, предварительную обработку данных, выбор спектральных диапазонов, обнаружение выбросов и определение числа латентных переменных и верификацию моделей.

Положения, выносимые на защиту. На защиту выносятся следующие положения:

- методические подходы разработки компьютерных моделей: регистрация цифровых сигналов, построение базы данных, создание наборов данных, предварительная обработка данных, выбор спектральных диапазонов, обнаружение выбросов и определение числа латентных переменных, верификация модели;
- модели классификации пищевых товаров на категории качества: зернового кофе по ботаническому виду, обжарке и географическому региону произрастания; гречневой крупы по времени урожая (весенний, осенний) и в зависимости от гидротермической обработки;
- компьютерные модели мониторинга качества пищевых товаров (плоды авокадо по наличию дефектов и зрелости, зерно гречихи по степени выполненности) с применением гиперспектральных изображений в видимом и ближнем инфракрасном диапазоне;
- калибровочные модели прогнозирования свойств и возможной фальсификации пищевых товаров (сливочное масло, гречневая крупа).

Степень достоверности и апробация результатов. Исследования проводились в течение 5 лет (2017-2022 г.г.). Достоверность результатов определялась методами математико-статистической и многомерной обработки с помощью программного обеспечения *Statistica Ver. 10, Unscrambler X 10.0.4, Origin 2018b, Matlab 2018b, Prediktera Breeze ver. 2021.1.0*. Основные результаты исследований обсуждены и доложены на VI Международной научно-практической конференции «Церевитиновские чтения – 2019» (ФГБОУ ВО «РЭУ им. Г.В. Плеханова», г. Москва, 22.03.2019 г.); международной научно-практической конференции «Качество, контроль и экспертиза на стоките» (Экономический университет – Варна, Болгария, 31.05.2019 г.); международной научно-практической конференции «Качество и сертификация на продукте» (Экономический университет – Варна, Болгария, 13.03.2020 г.); III Всероссийской научно-практической конференции с международным участием «Региональные рынки потребительских товаров: качество, экологичность, ответственность бизнеса» (Сибирский федеральный университет, Красноярск, 10.12.2020 г.); международной научно-практической конференции «XXXIV Международные плехановские чтения» (ФГБОУ ВО «РЭУ им. Г.В. Плеханова», г. Москва, 24.03.2021 г.); международной научно-практической конференции «Питание и здоровье: инновационные технологии» (ФГБОУ ВО «РЭУ им. Г.В. Плеханова», г. Москва, 31.03.2021 г.).

Публикации. По результатам диссертационной работы опубликовано 14 научных работ, включая 5 публикаций в журналах, входящих в перечень рецензируемых научных изданий Высшей Аттестационной Комиссии, из них 2 публикации категории К1 и 1 публикация категории К2, включенных в Перечень Высшей Аттестационной Комиссии при Минобрнауки Российской Федерации для публикации результатов диссертационных исследований, а также 2 публикации в журналах, индексируемых в наукометрической базе данных «Scopus» (Q3 и Q4).

Структура и объем работы. Диссертационная работа состоит из введения, обзора литературы, экспериментальной части, выводов, списка литературы и приложений. Содержание представлено на 166 страницах машинописного текста, включая 42 рисунка, 17 таблиц и 188 источников литературы, в том числе 160 зарубежных источников.

Глава 1 Мировые тенденции применения методов оптической спектроскопии и анализа изображений для идентификации и контроля качества пищевых продуктов

1.1 Теоретические аспекты методов оптической спектроскопии и методов анализа изображений

Методы оптической спектроскопии являются высокоточными неразрушающими аналитическими методами, которые позволяют проводить точные и быстрые измерения определенных свойств пищевых продуктов без предварительной обработки образцов.

Электромагнитный спектр делится на несколько областей в зависимости от энергии излучения, и в него входят: гамма-лучи, рентгеновские лучи, ультрафиолетовая область (УФ), видимая область спектра (*VIS*), инфракрасная область (ИК), микроволны и радиоволны.

Фото- и спектрофотометрические методы основаны на поглощении в видимой области спектра в диапазоне от 350 до 780 нм и применяются для количественного анализа определяемых веществ, компонента смеси или их окрашенных форм [20].

Инфракрасная область состоит из ближней (БИК), средней (ИК) и дальней инфракрасной области. ИК-спектроскопия успешно применяется для количественного анализа сложных смесей [91]. Это происходит главным образом потому, что в БИК-спектрах проявляются валентные и деформационные колебания химических связей с участием С-Н, N-H, S-H и О-Н. ИК-спектры возникают из более обширной спектральной информации об интенсивностях и частотах. ИК-спектроскопия обычно используется для определения химических функциональных групп в образцах пищевых продуктов, как в качественном, так и в количественном отношении.

Инфракрасная спектроскопия способна быстро представлять информацию об очень большом количестве анализируемых объектов, а полосы поглощения инфракрасных спектров чувствительны к физическим и химическим состояниям отдельных компонентов. БИК- и ИК-спектроскопия обладает более высокой интенсивностью сигнала по сравнению со спектроскопией комбинационного рассеяния, а спектроскопия средней инфракрасной области, в свою очередь, имеет преимущество по сравнению со спектроскопией ближней инфракрасной области, так как позволяет идентифицировать микроэлементы.

HSI рассматриваемое как многоспектральная визуализация с многочисленными полосами, получает данные в виде смежных спектральных полос [41,165]. Число диапазонов длин волн, захваченных системой спектральной визуализации, различно, обычно несколько сотен полос для *HSI* и около 5–20 длин волн для мультиспектральной визуализации. Благодаря интеграции основных функций визуализации и спектроскопии, спектральная визуализация может быть применена для количественного прогнозирования присущих химических и физических свойств, а также их пространственного распределения одновременно. В настоящее время большинство визуализирующих спектрометров предназначены для регистрации спектральной информации в видимой и ближней инфракрасной области (*VIS-NIR* от 400 до 2500 нм) [34].

Применение методов оптической спектроскопии и методов анализа изображений основано на информации, извлекаемой при эксперименте, которую необходимо анализировать с помощью специальных математических методов, т.е. с помощью хемометрики.

1.2 Хемометрические методы анализа спектральных данных

Хемометрика является самостоятельной поддисциплиной аналитической химии и преследует цели извлечения информации из исходных экспериментальных данных и последующее выявление наиболее важной аналитической информации. Она находится на стыке математики и химии, в ней применяются, в том числе, и статистические методы. Наиболее популярное и современное определение задач хемометрики в области химии было дано С. Волдом [179]:

- 1) Извлечение важной химической информации из химических данных.
- 2) Обработка и интерпретация полученной информации.
- 3) Получение данных, содержащих такую информацию.

Хемометрика позволяет заменить традиционные методы, характеризующиеся большими трудовыми затратами, использованием уникального оборудования, дорогостоящих реактивов, а также в некоторых случаях необходимостью работы с опасными веществами, на более быстрые и дешевые методы. Применение этих методов наиболее ярко выразилось в области ИК-спектроскопии, а первые работы в этой области решали задачи построения калибровочных моделей (градуировок), применяя метод главных компонент (МГК) и метод проекции на латентные структуры (ПЛС) [147].

Методы анализа в хемометрике разделяются на две группы, которые выполняют две основные задачи: 1) исследование данных с целью классификации и дискриминации; 2) прогнозирование новых значений. С одним блоком данных возможно выполнение задач первой группы, а методы второй группы, как минимум оперируют с двумя блоками данных (предикторы и отклики). К первой группе можно отнести такие методы, как кластерный анализ (КА) и дискриминантный анализ (ДА), метод опорных векторов, ко второй группе относятся МГК, проекция на латентные структуры, искус-

ственные нейронные сети и многие другие. Следует отметить, что методы, направленные на предсказание новых значений исследуемых объектов, получили большее развитие в сравнении с классификационными и дискриминационными методами, так как позволяют выполнять более углубленные аналитические задачи, такие как выявление фальсифицированной продукции или контроль качества товара. Общий вид методологии хемометрического моделирования представлен на рисунке 1.

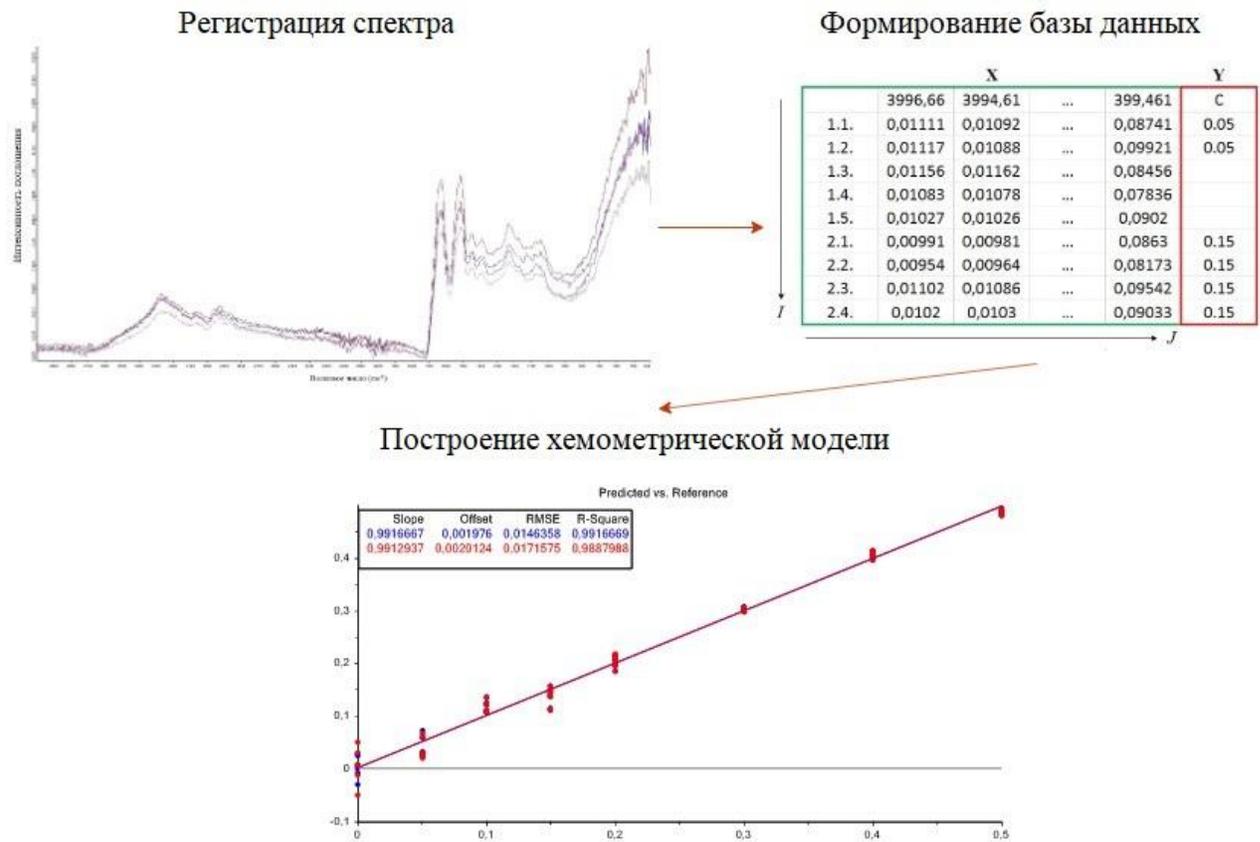


Рисунок 1 – Методология хемометрического моделирования

Источник: составлено автором по данным [27]

На данный момент приборы с легкостью позволяют проводить огромное количество измерений за небольшой промежуток времени. Несмотря на это, свойство мультиколлинеарности затрудняет извлечение необходимой информации из полученного массива данных. Чтобы изъять полезную информацию применяются математи-

ческие методы сжатия данных. При интерпретации исходных данных используют новые скрытые (латентные) переменные. При этом число переменных при сжатии должно значительно снижаться в сравнении с числом исходных переменных. Такое сжатие необходимо для представления информации в компактном виде, что делает визуализацию и интерпретацию более удобной и простой.

Наиболее распространенным методом сжатия данных является метод главных компонент, который заложил основы для создания более новых и точных методов, применяемых в хемометрике. Суть метода заключается в разложении исходной двумерной матрицы данных в виде произведения двух двумерных матриц T и P [27].

$$X = TP^t + E = \sum_{a=1}^A t_a p_a^t + E, \quad (1)$$

где T – матрица счетов (scores);

P – матрица нагрузок (loadings);

E – матрица остатков;

t_a – число столбцов в матрице T ;

p_a – число столбцов в матрице нагрузок P .

Число столбцов p_a в матрице нагрузок равняется эффективному рангу матрицы X . Эта величина обозначается как A , а также называется числом главных компонент (ГК), количество которых всегда меньше числа столбцов в матрице X .

МГК дает возможность решать задачи разделения, в таком случае он является, в том числе, и факторным анализом. Более того, этот метод применим к анализу любых химических данных. В зависимости от решаемых задач матрицы счетов T и нагрузок P могут восприниматься по-разному. Это означает, что число главных компонент в определенных случаях можно интерпретировать как фактическое количество отдельных химических соединений и их концентрации, а в других случаях, эти компоненты дают лишь общую информацию о разделении химических веществ в образце.

В решении задач количественного анализа [7] применяются два набора данных. Первый набор данных X – матрица аналитических сигналов, например, спектров; второй набор Y – матрица соответствующих химических показателей. С помощью этих

наборов данных проводится градуировка (калибровка), целью которой является построение математической модели, которая связывает наборы друг с другом. Такие модели в дальнейшем используются для предсказания значений y [120].

В решении задач количественного анализа используются градуировочные модели. Самый простой случай – это одномерная регрессия, соответствующая одному аналитическому сигналу [60]. Регрессионные методы анализа позволяют строить более сложные многомерные (множественные) регрессии, в которых задействуются несколько каналов [6]. Следует отметить, что в блоках данных X и Y могут содержаться погрешности, но современные регрессионные методы (РГК, ПЛС) и методы предварительной обработки данных позволяют работать при наличии погрешностей в обоих блоках данных [26].

Чтобы избавиться от погрешностей в спектральных данных и максимизировать различия в них, применяется предварительная обработка [148]. Достижение этой цели становится очень трудной задачей, если отношение сигнал/шум (S/N) полученного спектра очень мало [85]. Поскольку в ИК-спектроскопии используется свет низкой интенсивности, он обязательно будет вызывать менее чувствительные реакции в возбужденных молекулах [175]. Чтобы компенсировать это, уровни шума, обнаруженные прибором, должны быть сведены к минимуму, чтобы получить высокое отношение сигнал/шум [11]. Шум может быть вызван приборными ошибками во время сканирования, подготовкой образца, а также условиями окружающей среды, в которых генерируются спектры [176]. Одним из способов устранения шума является проведение нескольких сканов и их усреднение [96]; хотя для достижения значительных уровней шумоподавления может потребоваться большое количество повторных сканирований. Если это не приводит к адекватному снижению уровня шума, то используются методы сглаживания. Наиболее распространенные методы предварительной обработки включают нормализацию [122], производные [133], коррекцию мультипликативного рассеяния (MSC) света [77] и стандартная нормальная вариация (SNV). Ме-

тоды предварительной обработки, разработанные в последнее время, включают коррекцию ортогонального сигнала (*OSC*) [180], прямую коррекцию ортогонального сигнала (*DOSC*) [111] и коррекцию ортогонального вейвлета (*OWAVEC*) [67].

Конечная цель разработки калибровочных моделей – это точный и качественный прогноз неизвестных значений тестового набора образцов, поэтому оценка качества полученных моделей критически важна [78]. К набору образцов, в свою очередь, предъявляется важное правило – этот набор должен иметь реалистичное описание образцов и их свойств, которые могут встречаться и в будущих исследованиях [148].

Проверка достоверности прогностической способности калибровочной модели зависит от спектральных данных, содержащих информацию о компонентах, качественных свойствах и классификационной принадлежности образца, а также от измерений, которые проводятся традиционными стандартными методами [134]. Внутренняя проверка включает проверку осуществленной калибровки с использованием того же набора образцов, который использовался при разработке модели. Поэтому оценка, основанная на внутренней валидации, не совпадает с оценкой прогноза [121], которая включает валидацию калибровки с использованием независимого набора образцов, который не задействуется при разработке калибровочной модели. Обычно используются два метода проверки: независимая или внешняя проверка [65] и перекрестная проверка (кросс-валидация), в которой используется только набор данных калибровки. Внешняя проверка требует отдельного, большого и репрезентативного набора тестовых объектов, чтобы дать соответствующие и надежные оценки будущей способности модели к предсказанию [119].

Кросс-валидация является очень надежным методом проверки, который стремится проверить модель калибровки на независимом наборе данных испытаний, но, в отличие от внешних данных, он не использует образцы только для испытаний. Для перекрестной проверки последовательные образцы удаляются из набора калибровки. При частичной перекрестной проверке образцы удаляются группами; при полной пе-

рекрестной проверке все образцы удаляются по одному. После каждого удаления выполняется калибровка остальных образцов перед проверкой на удаленных образцах. Затем первый образец заменяется данными калибровки, а следующий образец удаляется. Процедура продолжается до тех пор, пока все образцы или группы образцов не будут удалены один раз [134].

Последний шаг в разработке калибровочной модели – это статистика, которая необходима для интерпретации отобранных данных и оценки эффективности и точности итоговой модели [177]. Статистические данные, наиболее часто используемые для количественного анализа хемометрических моделей, перечислены в таблице 1 и включают стандартную ошибку прогнозирования (*SEP*) или стандартную ошибку перекрестной проверки (*SECV*); уклон (*BIAS*); коэффициент детерминации (R^2) и отношение стандартного отклонения к стандартной ошибке прогнозирования (*RPD*) [148].

SEP измеряет, насколько хорошо калибровка будет выполняться в будущем анализе [68], и должна быть как можно ближе к стандартной лабораторной ошибке (*SEL*) для измеряемого компонента или свойства. *SEP* используется для оценки точности калибровочной модели путем указания изменчивости данных, полученных стандартными (эталонными) методами, от спектральных данных. Смещение (*BIAS*) показывает среднее значение, на которое отличаются результаты. Вместе *SEP* и смещение указывают и оценивают общую точность прогностической модели. *SEP* и смещение должны быть как можно ниже, с наклоном (*slope*), близким к 1,0 [177]. Когда выполняется кросс-валидация, точность калибровки указывается с помощью *SECV*. *RPD* позволяет оценивать *SEP* с точки зрения стандартного отклонения (*SD*) эталонных данных и дает представление об эффективности калибровочной модели [178]. Рекомендуемое значение *RPD* для применения спектроскопии при контроле качества составляет 5,0 или более [177]. Для целей скрининга будет достаточно значения более 3,1.

Таблица 1 – Статистические параметры оценки хемометрических моделей

Параметр	Равенство	Необходимое значение
<i>SD</i>	$\sqrt{\sum y^2 - \frac{ \sum y ^2}{n-1}}$	Минимальное значение
<i>SEL</i>	$\sqrt{\frac{\sum (y_1 - y_2)^2}{2n}}$	
<i>SEP/SECV</i>	$\sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i - BIAS)^2}{n-1}}$	Минимальное значение или близкое к SEL значение
<i>V-BIAS</i>	$\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)$	Как можно ближе к нулю
<i>R²</i>	$1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y}_i)^2}$	<0,90
<i>RPD</i>	$\frac{SD_{\hat{y}}}{SEP} \text{ или } \frac{SD_{\hat{y}}}{SECV}$	5,0 или выше

Источник: составлено автором по данным [65,148]

Коэффициент детерминации (R^2) указывает на количество вариаций в данных, которые объясняются уравнением калибровки. Значение R^2 должно быть высоким (<0,90), чтобы указывать на хорошую способность к прогнозированию, в то время как при низком значении (>0,64) невозможно получить стабильно высокую точность модели.

Точность результатов классификации также может быть выражена с точки зрения чувствительности и специфичности [66]. Чувствительность - это доля выборок, принадлежащих к определенной категории, правильно идентифицируемых моделью прогнозирования, соответствующей этому классу, т.е. это мера способности правильно прогнозировать истинные положительные результаты. Специфичность - это доля образцов, не принадлежащих к определенному классу, которые классифициру-

ются как инородные. Эти параметры являются ценными диагностическими инструментами для оценки классификации и подтверждения подлинности, поскольку классификационная модель должна не только правильно классифицировать образцы как принадлежащие к определенному классу или категории, но также отклонять образцы, не принадлежащие этому классу.

1.3 Применение методов оптической спектроскопии и методов анализа изображений для идентификации и контроля качества пищевых продуктов

Развитие методов оптической спектроскопии и методов анализа изображений в последние годы дает возможность получать уникальную информацию о белковых, углеводных и липидных структурах присущим пищевым продуктам. Во многих опубликованных работах была использована методология ИК-Фурье спектроскопии [1] для определения качества пищевых продуктов. Гиперспектральные методы применялись для классификации, градации и быстрой оценки качества пищевой продукции. Рассмотрим далее различные применения методов оптической спектроскопии в сочетании с хемометрикой в области оценки качества и выявления фальсификации отдельных видов пищевых товаров [12].

1.3.1 Продукты зерновых культур

Состав продукта является основой для определения пищевой ценности, поскольку химические компоненты продукта больше всего влияют на пищевую ценность. Основные компоненты с наибольшим процентным содержанием в составе зерновых и бобовых культур - это углеводы, за которыми следуют белки, влага, сырая клетчатка, зола и жир. Крахмал является разновидностью углеводов, включающий

спиральную амилозу и разветвленный амилопектин. Содержание амилозы широко варьируется в пределах одного вида, и были зарегистрированы следующие диапазоны амилозы: кукуруза 20–36%, сорго 21–35%, пшеница 17–29%, ячмень 11–26%, рис 8–37%, горох 34–37%, а картофель 18–23% [38]. Свойства крахмала, такие как соотношение амилозы и амилопектина, оказывают большое влияние на качество продуктов, таких как хлеб или лапша [89].

Белки являются основными структурными и функциональными компонентами зерномучных продуктов питания. Функциональные свойства белка, включая упругость и растяжимость, представляют собой те физико-химические свойства, которые влияют на качество и органолептические свойства основных пищевых продуктов. Влага как основной компонент представлена как растворитель, удерживаемый молекулярными и межмолекулярными связями, и как вода гидратации, удерживаемая водородной связью. Содержание воды в зерномучных продуктах питания во время хранения сильно влияет на качество хранения и успешность последующего прорастания. Клетчатка в таких продуктах в основном содержит целлюлозу, гемицеллюлозу, лигнин, пектин и камеди, однако не обеспечивают человека достаточным количеством энергии. Напротив, липиды являются наиболее концентрированной формой энергии, дающей вдвое больше энергии на грамм продукта, чем углеводы или белки.

Поскольку большинство из существующих методов обнаружения макронутриентов зерномучных продуктов питания являются разрушающими [50, 95], занимающими много времени и требуют длительного времени для подготовки образцов, поэтому применение новых экспресс-методов неразрушающего контроля крайне актуально.

Методы видимой инфракрасной спектроскопии обладают большим потенциалом для определения химических компонентов в зерномучных продуктах питания. В работе [137] показано, что *VIS-NIR*-спектроскопия в диапазоне от 400 до 1000 нм может быть использована в качестве неразрушающего метода для определения ядови-

того фитогемагглютинина в бобах путем контроля времени кипения стеблевой фасоли. Спектроскопические методы *NIR* и *IR* также применялись и сравнивались между собой в другом исследовании для определения химических свойств, относительно содержания влаги, белка, липидов и золы в 20 сортах сои. В исследовании [69] показано, что методы ближней и средней инфракрасной спектроскопии позволяют разработать достоверные модели ПЛС с высокими прогностическими способностями. Общее время, необходимое для подготовки и определения образцов, составило менее 5 минут по сравнению с 10–16 часами, необходимых при стандартных методах анализа. Тот же коллектив авторов [70] исследовал способность инфракрасной спектроскопии определять концентрацию влаги, белков, липидов, золы и углеводов бразильской сои. На основании регрессии проекции на латентные структуры наилучшие результаты были обнаружены для содержания белка ($R^2=0,81$) и влажности ($R^2=0,80$), за которым следовали липиды, пепел и углеводы.

Применение методов БИК-спектроскопии и хемометрических алгоритмов для анализа гречневой крупы показано в работе [84]. Авторами были разработаны точные хемометрические модели для прогноза содержания белка, крахмала и влажности. Наиболее интенсивные полосы в спектре в ближней инфракрасной области были связаны с наличием в образцах гречневой крупы белковых и жировых связей, фенольных соединений и влаги. В более поздних работах [170] рассмотрено использование ИК-спектроскопии и нейронных сетей для прогнозирования содержания белка и крахмала в гречке. На этапе предварительного анализа МГК применялся для выявления наиболее значимых главных компонент, которые в дальнейшем учитывались при построении моделей нейронных сетей. На основе используемых методов была достигнута высокая точность для прогноза крахмала ($R^2=0,90$) и менее высокая точность для прогноза белка ($R^2=0,75$).

В исследовании [108] был осуществлен точный прогноз содержания сырого белка в картофеле с помощью *NIR*-спектроскопии (1100–2300 нм). Метод ПЛС был

применен для разработки калибровочной модели с высокими коэффициентами детерминации в диапазоне от 0,86 до 0,95. Позднее в работе [156] был продемонстрирован потенциал гиперспектрального изображения, спектроскопии видимой и ближней инфракрасной областей для быстрых полевых измерений химического состава у двух видов клубней картофеля. В исследовании было показано, что *VIS-NIR*-спектроскопия (446-1125 нм) имеет наилучшую корреляцию для предсказания глюкозы, сахарозы и растворимых твердых веществ в измельченных образцах. Методы средней и ближней инфракрасной спектроскопии в сочетании с хемометрическими алгоритмами применялись для контроля химических и ферментативных соединений [167] в зерномучных пищевых продуктах, включая каротиноиды, флавоноиды, антоцианы, фенольные соединения, диосцин [103] и каталазу.

Гиперспектральные системы визуализации были применены для прогнозирования активности альфа-амилазы и содержания аминокислот, глюкозы, влаги и сахарозы в зерномучных продуктах питания. На основе метода ПЛС и полученной спектральной информации были предсказаны уровни активности фермента альфа-амилазы в ядрах пшеницы, причем самый высокий R^2 составил 0,88 [183]. Кроме того, для прогнозирования содержания аминокислот на основе концентраций сахарозы, глюкозы, фруктозы и азота в соевых бобах было исследовано в более ранних работах [131]. В этой работе была использована нелинейная регрессионная модель, основанная на преобразованном наборе спектральных данных второй производной. Была получена лучшая точность 0,74 для прогнозирования концентрации глюкозы, за которой следовали сахароза, азот и фруктоза с точностью 0,73, 0,61 и 0,60 соответственно, что указывало на потенциал гиперспектральной визуализации для оценки химического состава. Через несколько лет было проведено еще одно исследование по прогнозированию глюкозы, сахарозы, растворимых твердых веществ и других ингредиентов в картофеле на основе *HSI* (400–1000 нм) [156]. Глюкоза была предсказана с большей точностью 0,74, но прогноз содержания сахарозы и растворимого твердого вещества обладал слабой корреляцией с точностью 0,57 и 0,36 соответственно.

Содержание влаги является еще одним важным параметром, используемым для оценки качества основных продуктов питания. Коллективом авторов [88] были разработаны регрессионные модели для прогнозирования содержания влаги в соевых бобах в процессе сушки с помощью *HSI* (400–1000 нм). Лучший результат прогнозирования содержания влаги был достигнут с использованием автоматической сегментации и метода ПЛС с высоким значением R^2 , равным 0,94. Позднее содержание влаги в арахисе было определено с помощью ПЛС и продемонстрировало хорошие показатели при R^2 , равном 0,91, путем выбора шести специальных длин волн в 400–1000 нм *HSI* [93].

Сенсорные характеристики основных продуктов питания, включающие цвет, форму, жесткость, текстуру и внешние дефекты, являются важными компонентами качества пищи [150]. Образование визуального цвета происходит из-за отражения волн различной длины в видимой области света. Форма как общее свойство продукта определяется размером, весом или объемом и может влиять на предпочтения потребителей и их конечное потребление. Жесткость является одним из главных факторов, определяющих качество пищевых продуктов, отражающая текстуру и влажность пищевых продуктов. Кроме того, внешние дефекты, которые возникают либо в период роста, либо после сбора урожая, являются еще одним очень распространенным органолептическим признаком и серьезно влияют на качество пищи [56]. Поэтому точное и быстрое прогнозирование этих сенсорных признаков является важнейшим принципом в сельском хозяйстве и в пищевой промышленности.

Сенсорные характеристики, включая цвет и текстуру, были определены с помощью спектроскопии в видимой и ближней инфракрасной области. В исследовании, проведенном [126], спектроскопия *VIS-NIR* (400–2500 нм) эффективно использовалась для классификации неповрежденных сухих бобов путем прогнозирования качественных признаков консервирования, включая цвет и текстуру, а консервированные бобы

были классифицированы на две органолептические категории: «приемлемые» и «неприемлемые» с помощью линейной дискриминантной модели с точностью классификации около 72,60%.

Сенсорная оценка продуктов питания с использованием систем гиперспектрального изображения и визуализации в основном сфокусирована на определении цвета, жесткости и внешних дефектов. В частности, цвет свежих и сушеных соевых бобов был предсказан с помощью гиперспектральной визуализации (400–1000 нм) во время процесса сушки [88]. Также была разработана модель, основанная на методе проекции на латентные структуры, для определения цвета обработанных соевых бобов с удовлетворительным коэффициентом детерминации (R^2), равным 0,74.

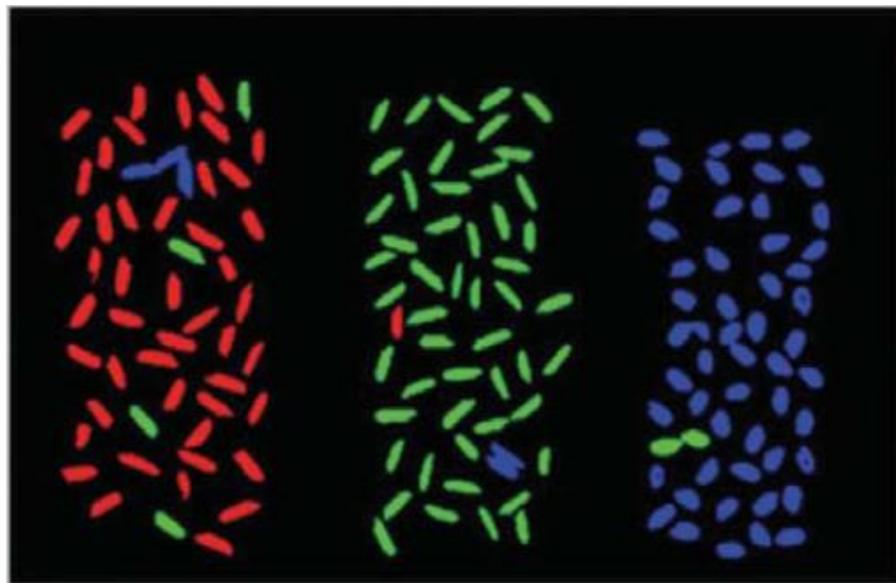


Рисунок 2 – Классификация риса, основанная на характеристиках цвета и формы, полученная методом гиперспектральной визуализации.

Источник: составлено автором по данным [172]

В недавнем исследовании [172] была разработана классификация образцов риса с помощью гиперспектральной визуализации в диапазоне 400–1000 нм на основе цвета и формы (рисунок 2). Образцы были проклассифицированы по пяти особенностям формы и по одному цветовому признаку. Далее было осуществлено построение

модели, основанной на алгоритме нейронной сети, а точность классификации составила - 94,45%. В работе [116] была построена модель определения твердости пшеницы с использованием регрессии на главные компоненты и проекции на латентные структуры с помощью гиперспектральной визуализации. Авторы отмечали, что использование модели ПЛС показывает лучшие результаты, чем применение метода РГК.

1.3.2 Животные и растительные масла

Исследования масел и жиров проводились многими отечественными и зарубежными авторами. Для оценки качества таких продуктов широкое распространение получили методы ИК-Фурье спектроскопии, с помощью которых возможна как классификация, так и количественный анализ компонентов в маслах.

Посредством ИК-Фурье спектроскопии до проведения математической обработки возможно назначение полос спектра различным функциональным группам образца. Авторы отмечали полосы поглощения при 1745 см^{-1} , 2853 см^{-1} , 2954 см^{-1} , 3005 см^{-1} , 966 см^{-1} , 3450 см^{-1} и 1640 см^{-1} , обусловленные поглощением карбонильной группы, валентных колебаний С-Н, =СН двойных связей и О-Н жиров, соответственно, как основные для масложировых продуктов [73]. Область около 1750 см^{-1} сложная, так как в ней проявляется поглощение алифатических сложных эфиров ($1750\text{-}1735\text{ см}^{-1}$), а также α , β -непредельные сложные эфиры ($1730\text{-}1715\text{ см}^{-1}$) [114]. Для различения растительных масел также важны слабые полосы при 722 и 3008 см^{-1} , описывающие поглощение деформационных и валентных колебаний =СН-групп. Отмечалось, что вид этих полос поглощения при купажировании взаимосвязан с механизмами конформационных изменений в композициях растительных масел [19].

При использовании хемометрических методов на основе измерений ИК-Фурье спектроскопии становится возможным проведение анализа компонентов и физико-химических свойств. Проведение традиционных контрольных анализов и дальнейшее

использование этих результатов при хемометрическом моделировании позволяет проводить анализ жирно-кислотного состава образцов, прогноз йодного числа и других технологических показателей на основании ИК-спектров.

В работе [161] были сопоставлены результаты газожидкостной хроматографии стандартов жирных кислот в соответствии с измерениями ИК-спектров сливочных, растительных масел и маргаринов. Указано, что различительной характеристикой растительных и сливочных масел также является содержание воды, которая сильно поглощает в области $3600-3000\text{ см}^{-1}$ и 1650 см^{-1} . С помощью МГК были выявлены основные спектральные отличия, а назначения ГК были присвоены концентрациям эфиров (ГК1), CH_2 и $\text{C}=\text{O}$ группировок (ГК2), степени ненасыщенности (ГК4). Было показано, что ИК-спектроскопия в сочетании с математическими методами позволяет извлекать достоверную информацию и проводить быструю классификацию жировых пищевых продуктов.

Использование ИК-Фурье и Рамановской спектроскопии с хемометрическими алгоритмами для проверки качества растительных масел было проведено в работе [162]. В качестве контрольных испытаний проводилось измерение йодного числа и газовая хроматография метиловых эфиров исходных образцов масел. Авторы отмечали, что отношение интенсивности полос поглощения 3005 и 2856 см^{-1} может служить мерой степени насыщенности масел. Разработанные модели ПЛС для прогнозирования значений йодного числа отличались высокой точностью при использовании как ИК-спектров ($R^2=0,93$), так и спектров комбинационного рассеяния ($R^2=0,97$). Модель прогнозирования степени ненасыщенности, основанная на Рамановских спектрах, отличалась высокой точностью ($R^2=0,93$), тогда как аналогичная модель на ИК-спектрах должна разделяться на отдельные подобласти для адекватного прогноза.

Решение задач выявления фальсифицированной масложировой продукции осуществляется путем построения калибровочной модели из смесей образцов в различных концентрациях. Применение хемометрических алгоритмов для выявления фальсификации сливочного масла говяжьим жиром было представлено в работе [141]. В

исследовании было показано, что пики поглощения при 1117 см^{-1} (валентные колебания —СН в жирных кислотах) и при 1097 см^{-1} (деформационные колебания —СН в жирных кислотах) в говяжьем жире имеют гораздо более высокую интенсивность в сравнении со сливочным маслом. Поэтому построение модели ПЛС проводилось в спектральном диапазоне $1500\text{--}1000\text{ см}^{-1}$, что обеспечило высокую точность модели ($R^2=0,996$). Аналогичным путем были разработаны модели для выявления фальсификации сливочного масла пальмовым маслом [57] ($R^2=0,977$) и куриным жиром [135] ($R^2=0,981$).

В последних исследованиях была также показана возможность осуществления классификации [125] и выявления фальсификации оливкового масла. В работе [159] на основе измерений ИК-спектров смесей пальмового масла с оливковым и маслом семян тыквы были разработаны модели ПЛС и РГК для количественного анализа, а для классификации. Особое внимание при количественном анализе авторы уделили выбору спектрального диапазона и предварительной обработке спектральных данных, что привело к созданию 36 моделей. Наиболее точной оказалась модель ПЛС в спектральных диапазонах $3100\text{--}2750$ и $1500\text{--}663\text{ см}^{-1}$ по первой производной спектров ($R^2=0,996$).

Для прогнозирования содержания основных жирных кислот в оливковых маслах было отобрано 216 образцов с различных плантаций в Австралии и разработана калибровка на основе измерений БИК-спектроскопии [117]. Коэффициенты детерминации (R^2) были относительно низкими в моделях определения незначительных жирных кислот, таких как стеариновая (0,86) и линоленовая кислота (0,85), а для определения основных жирных кислот, т.е. олеиновой (0,99) и линолевой (1,00), обеспечен высокий уровень точности сравнимый с эталонными методами. Определение содержания жирных кислот в оливковых маслах с помощью ИК-Фурье спектроскопии было представлено в работе [114]. Значения коэффициентов детерминации для всех моделей определения жирных кислот были в диапазоне 0,96-0,99, а предел обнаружения составил для олеиновой кислоты 3,0%, линолевой кислоты 0,5%, общего содержания

насыщенных 1,3%, мононенасыщенных 0,3% и полиненасыщенных 0,3% кислот. В исследовании [168] с помощью ИК-Фурье спектроскопии проведена оценка качественных параметров оливкового масла, а именно окислительной стабильности, цветных пигментов, профиль жирных кислот и фенольный состав.

В нескольких исследованиях хемометрика на основе спектроскопических данных применялась для разработки моделей прогнозирования содержания фенольных соединений [44] и токоферолов в оливковом масле [136]. Количественный анализ общего содержания фенольных соединений на основе ИК-спектроскопии в нерафинированном оливковом масле в была представлена в работе [82], а содержание отдельных основных фенольных соединений на основе БИК-спектроскопии с помощью ПЛС в [37].

Быстрое определение общего содержания фенольных соединений и антиоксидантной активности оливковых масел на основании ИК-Фурье спектроскопии представлено в [46]. Показано, что хорошие результаты при оценке содержания фенольных соединений могут быть достигнуты при использовании диапазона $3610-816\text{ см}^{-1}$, тогда как диапазон $3707-1105\text{ см}^{-1}$ не применим для оценки антиоксидантной активности.

Применение БИК-спектроскопии для оценки растительных масел также показано в исследованиях [51,75]. В работе [75] исследована возможность прогноза кислотности отходов кулинарных масел на основе стандартного анализа и измерений БИК-спектров. Продемонстрированы калибровочные модели ПЛС: в полном спектральном интервале $800-2200\text{ нм}$ ($R^2=0,970$); в интервале $950-1650\text{ нм}$, то есть в типичном интервале для переносных БИК-спектрометров ($R^2=0,905$). Снижение точности обусловлено отсутствием информации, относящейся к первому обертому колебаний C–H нескольких химических групп ($=\text{CH}-$, $-\text{CH}_3$, $-\text{CH}_2-$), характерному для жирных кислот растительных масел и расположенному при 1720 нм . Исследование [51] было направлено на прогнозирование содержания жира в измельченных оливках (г) и в сыром продукте (%) с использованием *UV-VIS-NIR*-спектроскопии, ЯМР и экстракцией

жира по Сокслету. В результате была подтверждена методологическая ошибка разработки спектроскопических моделей ПЛС для непосредственного измерения содержания жира (г) в плодах, так как спектроскопия чувствительна к концентрациям, а не к общему количеству соединений одного класса (г). С другой стороны, модель ПЛС для прогнозирования содержания жира в сыром продукте (%) обладает высокой точностью ($R^2=0,93$).

Быстрая идентификация подлинности и выявление фальсификации масла семян черного тмина холодного отжима представлено в работе [31]. Методами многомерного анализа разработаны модели для различения рафинированного от масла холодного отжима, для выявления фальсификации масла семян черного тмина соевым маслом – на основе ИК-спектроскопии; классификационная модель *SIMCA* для различения соевого, хлопкового, нерафинированного оливкового масел и масел грецкого ореха и виноградной косточки – на основе флуоресцентной спектроскопии.

1.3.3 Плодоовощная продукция

Развитие экономики и улучшение уровня жизни повлекло за собой значительный прирост потребления свежих плодов и овощей. С точки зрения потребителя от таких продуктов ожидаются высокие качественные характеристики, такие как вкус, зрелость, плотность, кислотность и т.д. Однако, параметры качества плодоовощных продуктов контролируются с использованием традиционных (стандартных) методов, которые могут являться субъективными и отнимают много времени [170]. Для улучшения процесса контроля качества плодов и овощей был исследован потенциал различных технологий, среди которых методы оптической спектроскопии и методы анализа изображений привлекают особое внимание из-за следующих преимуществ:

- 1) Эти методы являются неразрушающими, позволяют получать информацию о внутренних параметрах плодов и овощей без повреждения их поверхности;

2) Измерительный процесс простой и быстрый, не требует сложных предварительных обработок или химических реакций;

3) Методы позволяют обнаруживать одновременно комплекс показателей качества.

Спектры видимой и ближней инфракрасной области содержат обширную информацию о колебаниях О-Н, С-Н и N-H связей, что позволяет оценивать различные качественные параметры плодов и овощей. Обзор диапазонов волн, содержащих типичные полосы поглощения для определенных химических групп представлен в таблице 2.

Таблица 2 – Диапазоны волн, содержащие типичные полосы поглощения для определенных химических групп плодоовощных продуктов

Параметр качества	Химическая группа	Длина волны, нм	Ссылка
Сахар	О-Н	1190,1400	[97]
Содержание сухих растворимых веществ	С-Н	910	[146]
	О-Н	960,1450	[106]
	С-Н, О-Н	1210	
	О-Н	975	[164]
	О-Н	960,1180, 1450, 2000	
	С-Н	963	
	Комбинационные полосы С-Н и О-Н	2000-2500	[187]
	С-Н, О-Н	950-1075	[186]
Кислотность	С-Н из СООН	1607	[105]
	О-Н карбоксильных кислот	1127	

Параметр качества	Химическая группа	Длина волны, нм	Ссылка
	С=О насыщенных и ненасыщенных карбоксильных кислот	1437	
рН	С-Н	768	[79]
	О-Н	986	

Источник: составлено автором по данным [79,97,105,106,146,164,186,187]

На основе методов оптической спектроскопии и методов многомерного анализа были исследованы многие параметры плодов и овощей, такие как содержание сухого вещества [124], твердость плодов [110], водородного показателя, содержания сахара и сухих растворимых веществ [106,146,164,187].

В работе [130] с высокой точностью ($R^2=0,94$) был осуществлен прогноз кислотности образцов киви по спектрам пропускания в диапазоне 400-1000 нм. При разработке модели были использованы методы предварительной обработки спектров, такие как *SNV*-коррекция, центрирование и первая производная. При использовании *VIS-NIR*-спектроскопии (408-2492 нм) [104] была достигнута более высокая точность ($R^2=0,95$) и качество модели ($SEP=0,28\%$).

Методами гиперспектральной визуализации были исследованы внешние повреждения плодов и овощей. Классификация клубники [135] в зависимости от различного уровня внешнего механического воздействия была получена с помощью гиперспектральной визуализации в диапазоне 650-1000 нм. Для решения задачи был проведен сравнительный анализ использования различных дискриминационных методов, включая линейный дискриминантный анализ и искусственные нейронные сети. Было обнаружено, что обнаруженные повреждения значительно уменьшаются в течение срока хранения. В исследовании [163] была изучена взаимосвязь спектров ближней

инфракрасной области и цветовых характеристик клубники в колориметрической системе $CIEL^*a^*b^*$.

В работе [48] были разработаны модели прогнозирования содержания ликопина, твердости, цветовых характеристик свежих томатов на основе измерений $VIS-NIR$ -спектров. Авторами были предложены модели предсказания содержания ликопина в томатах (мг/кг) по спектральным данным видимой области 410-750 нм ($R^2=0,98$) и по значению $Hunter$ ($R^2=0,94$). При построении моделей использовалось сглаживание Савицкого-Голая. Дополнительно, в используемом диапазоне видимой области были рассмотрены взаимосвязи полос поглощения свежих томатов и ликопина, экстрагированного хлороформом.

Использование БИК-спектроскопии и хеометрических методов для прогнозирования содержания сухого вещества, сухих растворимых веществ ($^{\circ}Brix$), жесткости кожуры и мякоти томатов в зависимости от степени созревания рассмотрено в исследовании [157]. Показано, что в процессе созревания уменьшается твердость как мякоти, так и кожуры томатов. Прогнозируемые хеометрическими моделями значения параметров не имели значительных различий со значениями, полученными эталонными методами.

Прогноз содержания витамина С в апельсинах был осуществлен с использованием спектрального диапазона 833-2500 нм [182]. Решение данной задачи осуществлялось путем последовательного моделирования с применением многочисленных алгоритмов предварительной обработки спектральных данных. В процессе разработки модели был найден оптимальный спектральный диапазон 1333-1835 нм, построенная в этом диапазоне модель ПЛС отличалась высокой точностью предсказания витамина С ($R^2=0,96$).

В исследовании [181] была разработана модель классификации различных видов персиков на основе спектров отражения в области 401-1000 нм и метода главных компонент. Модель включала 8 главных компонент, анализ которых позволил разработать модель дискриминантного анализа с точностью 100%. В другой работе [138]

сравнивалась точность различных методов дискриминации для классификации персиков. Показано, что модели дискриминантного анализа и *SIMCA* имели точность более 92%, а модель *DPLS* более высокую точность 95%.

Классификация различных сортов клубники на основе спектроскопии в диапазоне от 1100 до 2200 нм представлена в работе [184]. Была оценена результативность модели опорных векторов, нейронных сетей и дискриминантного анализа, наилучшие результаты были получены с помощью модели нейронной сети с точностью 97,14%. Впоследствии была предпринята попытка упростить дискриминацию [73]. Спектры трех сортов клубники в диапазоне 350-2500 нм были подвергнуты предварительной обработке: вторая производная, *SNV*-коррекция и метод скользящего среднего. В результате моделирования была получена модель *PLS-ANN* ($R^2=0,97$).

В исследовании [59] была рассмотрена возможность *VIS-NIR*-спектроскопии отражения для измерения общего содержания каротиноидов, α -каротина, β -каротина и лютеина в бананах. Модели ПЛС были основаны на спектральных данных, преобразованных с помощью сглаживания Савицкого-Голая и первой производной. Показано, что прогнозирование общего содержания каротиноидов и β -каротина возможно с высокой точностью ($R^2=0,98$, $0,96$ соответственно), хороший результат получен для определения α -каротина ($R^2=0,91$). Однако, модель предсказания содержания лютеина в бананах имела неудовлетворительные результаты ($R^2=0,75$). С учетом того, что α - и β -каротины составляют 90% от общего содержания каротиноидов в бананах, измерение содержания каротиноидов в бананах осуществимо методами *VIS-NIR*-спектроскопии.

Таким образом, мировой опыт в применении методов оптической спектроскопии и методов анализа изображений для контроля качества отдельных групп пищевых товаров, а именно: зерновых культур, животных и растительных масел, плодоовощной продукции, показывает перспективность использования данных методов для идентификации и контроля качества пищевых продуктов.

Глава 2 Объекты и методы исследования

2.1 Постановка научной задачи

Идентификация пищевой продукции – это процедура установления соответствия конкретной продукции образцу или ее описанию. Для пищевой продукции такая процедура проводится по характерным признакам или критериям идентификации. В соответствии с ТР ТС 021/2011 идентификацию пищевой продукции проводят по наименованию, а также визуальными, органолептическими и аналитическими методами. Органолептические методы могут быть субъективными, а аналитические методы требуют высокой квалификации, трудоемки и зачастую несут разрушающий характер.

Контроль качества пищевой продукции, то есть проверка соответствия ее количественных и качественных характеристик установленным требованиям, осуществляется аналогичными методами в соответствии с требованиями нормативно-технической документации. Включение отечественной системы стандартизации в ИСО предъявляет требования к процедуре контроля качества, а именно необходимость мониторинга качества пищевой продукции на всех этапах жизненного цикла.

Для оперативной работы товароведов, таможенных экспертов и лабораторий по качеству пищевой продукции в торговых сетях необходимо получать информацию о качестве продукции в режиме реального времени. Чтобы обеспечить эффективность контроля качества пищевой продукции в последнее время используют неразрушающие методы: машинное зрение, системы принятия решения с использованием искусственного интеллекта, недорогостоящие оптические и спектральные датчики, гиперспектральное изображение. Такие методы способны анализировать качество пищевой продукции в реальном времени и на всех этапах жизненного цикла, что является их преимуществом в сравнении со стандартными органолептическими и аналитическими методами.

Пищевые продукты как объекты сложного химического состава, обладают разным качеством, что предъявляет дополнительные требования к методам идентификации и контроля их качества. Многие дефекты пищевых продуктов не являются явными, а для их контроля требуются различные, в том числе разрушающие, методы (пенетрометрия, хроматография, определение влажности и др.). Поэтому разработка методических подходов применения оптической спектроскопии и гиперспектрального изображения, как быстрых и неразрушающих методов, для идентификации и контроля качества пищевых продуктов является актуальной.

Применение методов оптической спектроскопии и гиперспектрального изображения в решении задач идентификации и контроля качества пищевой продукции связано с обработкой полученных данных и построением компьютерных моделей с помощью многомерных методов. Использование многомерных методов в решении задач классификации и прогнозирования свойств рассмотрено в работах отечественных ученых (О.Е.Родионова, А.Л.Померанцев, Ю.Б.Монахова, А.Ю.Богомолов). Данные работы посвящены решению задач в области аналитической химии, а работы, посвященные решению задач идентификации и контроля пищевых продуктов, немногочисленны и не получили достаточного внимания.

В настоящей работе предложены методические подходы применения оптической спектроскопии и гиперспектрального изображения для идентификации и контроля качества пищевых продуктов. В зависимости от объекта контроля и решаемой задачи предложено использование различных инструментальных методов (спектроколориметрия, ИК-Фурье спектроскопия, UV-VIS-NIR-спектроскопия, гиперспектральное изображение в видимом и ближнем инфракрасном диапазонах) и многомерных методов (методы количественного анализа и классификации). Описаны методы предварительной обработки спектральных данных и анализа гиперспектральных изображений.

Выработаны практические рекомендации по применению инструментальных методов и многомерного моделирования для классификации и количественного анализа

для различных групп пищевых товаров. Применение гиперспектрального изображения для контроля качества пищевых продуктов позволяет одновременно использовать как химическую, так и пространственную информацию об объекте, что увеличивает вариативность возможных приложений. В данной работе отработаны методические подходы применения гиперспектральных изображений для разработки компьютерных моделей идентификации пищевой продукции: классификация плодов авокадо Хасс на категории качества по наличию дефектов; определение влажности и сухого вещества в составе плодов авокадо; сортировка авокадо по сроку хранения с использованием вегетационных индексов и визуализации гиперспектральных изображений по их значениям; градация зерна гречихи по степени выполненности. Продемонстрирована возможность неразрушающей оценки фруктов с толстой кожурой, включая состояние мякоти и внутренних дефектов, с применением гиперспектрального изображения в видимой области на примере плодов авокадо Хасс.

Например, использование UV-VIS-NIR-спектроскопии позволяет, с одной стороны, определить по VIS-диапазону цветовые характеристики и центры окраски, с другой — идентифицировать по NIR-диапазону функциональные группы компонентов химического состава при оценке объектов растительного происхождения и показано в работе для градации гречневой крупы, различающейся по размеру ядра, времени урожая, сроку хранения, в зависимости от ГТО.

В свою очередь, при контроле качества сливочного масла рекомендуется использование ИК-Фурье спектроскопии, так как проявляемые в ИК-диапазоне колебания химических связей позволяют точно отличать различные виды масел и их смесей. ИК-Фурье спектроскопия также является оптимальным инструментальным методом для идентификации зернового кофе по ботаническому виду, географическому месту произрастания и обжарке. Несмотря на то, что ботанический вид и обжарка определяются органолептически, сортировка измельченного зернового кофе по таким идентификационным признакам является затруднительной, особенно при контроле «микс» кофе с добавлением Арабики и Робусты в различных пропорциях.

Предлагаются методы предварительной обработки спектральных данных в зависимости от задачи и объекта исследования: отбор спектральных диапазонов, содержащих значимую информацию по решаемой задаче, с помощью МГК или ANOVA; коррекция мультипликативного рассеяния (MSC) света; стандартная нормальная вариация (SNV); преобразование спектров отражения в спектры поглощения по формуле $A = \text{Log}1/R$; сглаживание спектральных данных методами скользящего среднего, Савицкого-Голея и др.

Предложенные в настоящей работе методические подходы применения оптической спектроскопии и гиперспектрального изображения для идентификации и контроля качества пищевых продуктов позволяют: проводить быстрый и неразрушающий мониторинг качества пищевой продукции на всех этапах жизненного цикла; применять компьютерные модели идентификации пищевой продукции, дополняя стандартные методики (ГОСТ 31700-2012, UNECE STANDARD FFV-42:2019).

Отработка методических подходов проведена на различных объектах пищевых продуктов, отличающихся по показателям качества и идентификационным признакам. Использование ИК- и UV-VIS-NIR-спектроскопии позволяет оценивать качество поверхности объекта, а за счет глубины проникновения по различиям функциональных групп компонентов состава выборки пищевой продукции – о физиологических процессах и различиях между образцами. Применение методов многомерного моделирования используется для извлечения наиболее статистически значимой информации из массива спектральных данных для разработки компьютерных моделей идентификации и прогнозирования свойств пищевой продукции. Комплексная схема методических подходов исследования представлена на рисунке 3.

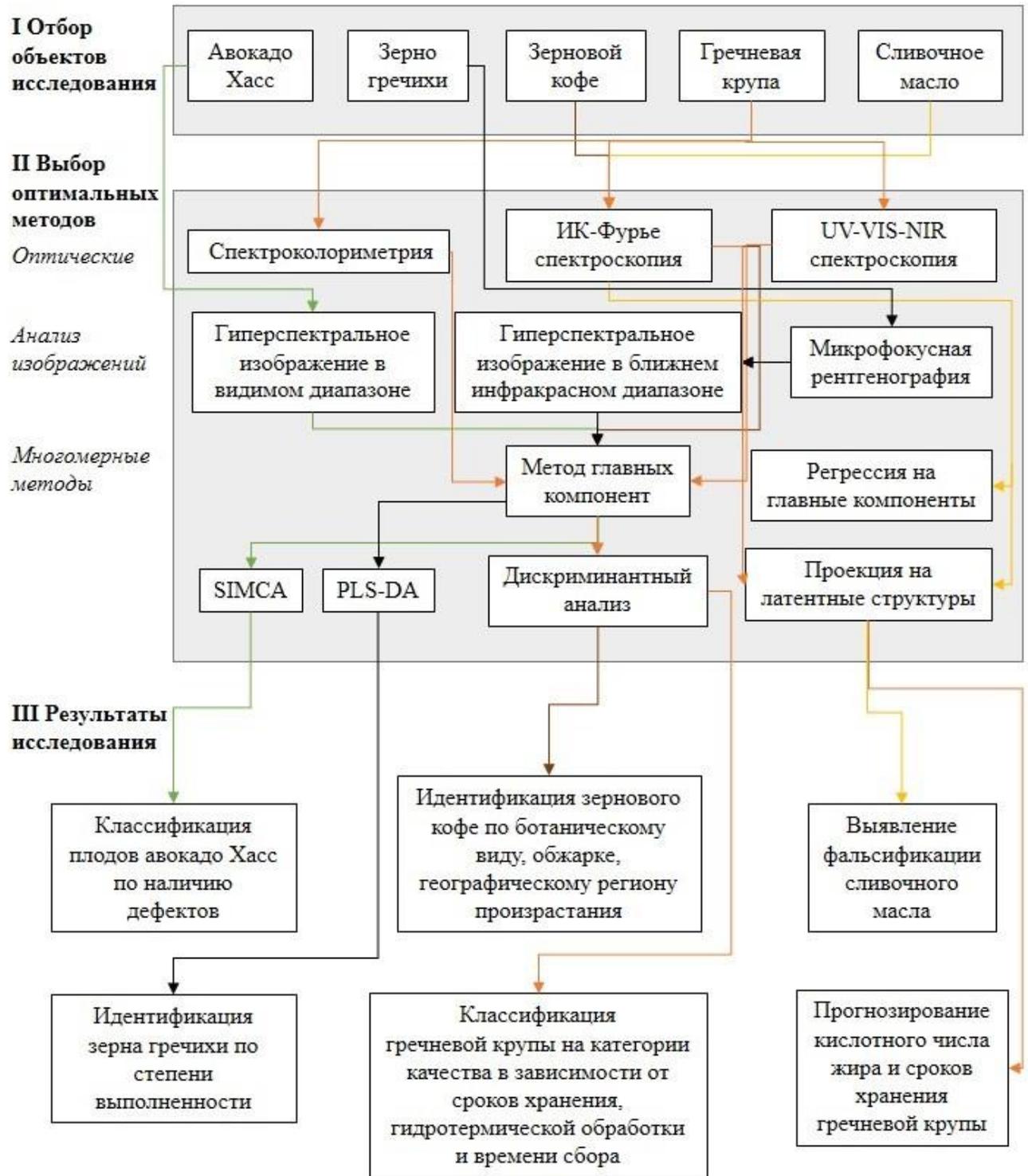


Рисунок 3 – Комплексная схема методических подходов применения оптической спектроскопии и гиперспектрального изображения для идентификации и контроля качества пищевых продуктов

Источник: составлено автором

2.2 Характеристика объектов исследования

Экспериментальное исследование проводилось в несколько этапов. На первом этапе проводился отбор и обоснование исследуемых образцов пищевых товаров. В качестве объектов исследования были отобраны отдельные группы пищевых товаров, а именно: сливочное масло, гречневая крупа и зерно гречихи, зерновой кофе, плоды авокадо сорта Хасс. Обоснование выбора объектов исследований:

1) Сливочное масло выбрано в качестве объекта исследования в связи с проблемами фальсификации, которые возникают из-за высокой себестоимости продуктов с содержанием молочного жира. Недобросовестные производители как в Российской Федерации, так и за рубежом, прибегают к фальсификации или подмене компонентов состава продуктов, содержащих молочный жир.

2) Контроль качества гречневой крупы и зерна гречихи согласно нормативной документации осуществляется традиционными химическими методами. Современные методы сортировки, включая использование фотосепараторов в видимом диапазоне, несовершенны. Поэтому целесообразна как разработка быстрых методов контроля качества с использованием спектроскопии и гиперспектральных изображений, так и получение данных, для улучшения контроля при использовании фотосепараторов.

3) Для контроля качества зернового кофе используются методы газовой и жидкостной хроматографии, которые являются точными и дают подробную информацию из-за разделения различных соединений. Но эти методы анализа трудоемки, сложны по пробоподготовке и не пригодны для оперативного мониторинга качества. Поэтому в исследовании проведена разработка и апробация метода идентификации зернового кофе по ботаническому виду, географическому месту произрастания и обжарке с применением методов ИК-Фурье спектроскопии и многомерного анализа.

4) Современные методы идентификации плодов авокадо носят разрушающий характер. Такие методы как измерение жесткости плодов с помощью пенетрометра или определение сухого остатка являются точными, но изымают исследуемые плоды из логистической цепочки. Поэтому в данной работе осуществлена разработка неразрушающей системы идентификации плодов авокадо на категории качества по наличию дефектов методами *HSI* и хеометрики.

Характеристика исследуемых образцов пищевых товаров представлена в таблице 3.

Таблица 3 – Характеристика исследуемых образцов пищевых товаров

Раздел	Наименование образца	Количество образцов, <i>n</i>	Описание
3.1	Гречневая крупа	20	Образцы пропаренной и непропаренной гречневой крупы, отличающиеся по срокам хранения. Образцы крупы под номерами 1-4,10,18 – непропаренные, все остальные номера присвоены образцам пропаренной крупы. Среди образцов непропаренной гречневой крупы, образцы 1 и 2 имеют срок хранения около одного года [24].
3.2	Гречневая крупа	32	Два набора образцов гречневой крупы: 1) образцы пропаренной и непропаренной гречневой крупы, отличающиеся по срокам хранения ($n=20$), указанные в п. 3.1.; 2) образцы крупы, отличающиеся по времени урожая (осень/весна, 2019/2020 г.) и размеру ядер ($n=12$). Образцы пропаренной гречневой крупы осеннего и весеннего урожая из второго набора представлены в 6 фракциях по размеру ядра [23].
3.3	Зерновой кофе	50	Образцы кофе, различающиеся по ботаническому виду – Арабика и Робуста, по географическому месту произрастания – Азия (Индия, Мьянма, Вьетнам), Америка (Бразилия, Перу, Колумбия, Сальвадор), Африка (Танзания, Уганда, Эфиопия), по обжарке – не обжаренное и обжаренное [128].
4.1	Сливочное масло	Калибровочный набор – 7; тестовый набор – 13.	В качестве образца-эталоны сливочного масла с гарантированным составом использовалось сливочное масло «Рузское» (м.д.м.ж. 82,5%). Этот образец был принят за эталон в соответствии с данными исследования [8]. Для подготовки калибровочного набора образцов был использован образец пальмового масла – масло пальмовое рафинированное отбеленное дезодорированное («DUKE'S», Малайзия). Всего было приготовлено 7

Раздел	Наименование образца	Количество образцов, <i>n</i>	Описание
			смесей, в концентрациях от 5 до 50%. Смеси сливочного и пальмового масел заданного состава получали при плавлении на водяной бане, при температуре 57 °С. Тестовый набор составлен из образцов сливочного и растительно-сливочного масла, приобретенных в розничных торговых сетях России и Болгарии.
4.2	Гречневая крупа	7	Калибровочный набор образцов гречневой крупы, отличающихся по срокам хранения (от 1 до 15 месяцев).
5.1	Зерно гречихи	81	Партия зерна гречихи (производство ООО «Бийский элеватор», Алтайский край), включающая в себя недоброкачественные зерна. Исследуемая партия зерна не была подвергнута обрушиванию и гидротермической обработке. Из партии было отобрано 81 зерно для проведения исследования.
5.2	Плоды авокадо сорта Хасс	10	Образцы плодов авокадо сорта Хасс (страна происхождения: Танзания; урожайность 2021 г.), которые были отобраны из одной партии и использованы для разработки хеометрической модели классификации по классам качества в зависимости от наличия дефектов. Плоды авокадо хранились в течение 10 дней (срок хранения, установленный производителем, составляет 14 дней) [129].
5.3	Плоды авокадо сорта Хасс	Калибровочный набор -9; тестовый набор – 10.	В качестве калибровочного набора отобраны плоды авокадо из одной партии сорта Хасс пурпурного и темно-зеленого цвета (страна происхождения: Танзания; урожай 2021 года). В тестовый набор отобраны плоды авокадо сорта Хасс багрового и темно-зеленого цвета [17].
5.4	Плоды авокадо сорта Хасс	47	Плоды авокадо сорта Хасс пурпурного и темно-зеленого цвета (страна происхождения: Танзания; урожай 2021 года). Образцы хранились в темном месте при комнатной температуре в течение 12 дней. Для данных 40 образцов в течение сроков хранения образцов отмечалось их потемнение кожицы и снижение плотности мезокарпия, при этом указанные процессы ускорились на 7 день хранения [16].

Источник: составлено автором

Второй этап был посвящен исследованию качества сливочных масел. На этом этапе также происходил анализ научной и технической литературы по применению как инструментальных, так и хеометрических методов исследования.

На третьем этапе проводился анализ спектральных данных и гиперспектральных изображений исследованных пищевых товаров с помощью хемометрических методов «без обучения»: МГК и КА. Указанные хемометрические методы были апробированы на следующих объектах исследования: зерновой кофе, гречневая крупа, плоды авокадо.

Четвертый этап исследования включал разработку классификационных моделей идентификации пищевых товаров методами «с обучением»: дискриминантный анализ, *PCA-DA* и *PLS-DA* – модификации дискриминантного анализа. На четвертом этапе были разработаны хемометрические модели определения страны происхождения зернового кофе, градации образцов гречневой крупы на категории качества в зависимости от сроков хранения, гидротермической обработки и времени сбора урожая.

На пятом этапе исследования проводилась идентификация пищевых товаров методами гиперспектрального изображения. С помощью *HSI* и хемометрических методов (анализ главных компонент, *PLS-DA*, *SIMCA*) получены следующие системы классификации: сортировка зерна гречихи по степени выполненности на основе гиперспектральных изображений в БИК-диапазоне от 1000 до 1700 нм; классификация плодов авокадо Хасс по наличию дефектов на основе гиперспектральных изображений в *VIS*-диапазоне от 400 до 1000 нм.

2.3 Методы проведения исследований

При выполнении диссертационного исследования отдельные группы пищевых товаров исследовали с помощью инструментальных методов и методов многомерного анализа. В качестве инструментальных методов применяли спектроскопические и спектроколориметрические методы, методы гиперспектрального изображения, а также методы контрольного анализа, указанные в действующей нормативной и технической документации.

2.3.1 Инструментальные методы

Спектроскопические методы. Одним из использованных спектроскопических методов была ИК-Фурье спектроскопия. ИК-спектры образцов регистрировали с помощью ИК-Фурье спектрометра *Bruker ALPHA* (*Bruker*, Германия) с алмазным модулем НПВО в средней инфракрасной области спектра $4000\text{--}400\text{ см}^{-1}$ с разрешающей способностью спектрометра 2 см^{-1} (рис. 4). Вне зависимости от исследуемого объекта измерение проводилось минимум в 5 повторениях для создания репрезентативной базы данных спектральных измерений. До и после регистрации спектров кристалл спектрометра очищали медицинским этиловым спиртом. С помощью встроенного программного обеспечения (*OPUS 7.3.5*) из каждого спектра автоматически вычитались фоновые шумы, и проводилась нормализация спектральных данных. После регистрации спектральные данные экспортируются из встроенного программного обеспечения в *Excel* в бинарном формате *dpt*, в котором формируется матрица данных для дальнейшего анализа.

Преимущества ИК-Фурье спектроскопии заключаются в простоте и скорости проведения измерений, возможности исследования на месте («*in situ*»), а также значительным количеством химической информации, содержащейся в ИК-спектрах.



Рисунок 4 – ИК-Фурье спектрометр Bruker ALPHA

Источник: составлено автором

ИК-Фурье спектроскопия применялась при исследовании зернового кофе (п. 3.3) сливочных масел (п. 4.1) и гречневой крупы (п. 4.2).

При измерении ИК-спектров зернового кофе для избавления от фоновых шумов, возникающих при измерении ИК-спектров грубодисперсных объектов, образцы зернового кофе измельчали в лабораторной мельнице с последующим просеиванием через сито с диаметром отверстий, равным 0,43 мм, таким образом размер частиц составлял от 0,35 мм до 0,42 мм.

ИК-спектроскопия также использована при исследовании сливочного масла (тестовый набор), а также его смесей с пальмовым маслом (калибровочный набор) и дальнейшей разработке калибровочной модели. Регистрацию спектров образцов из калибровочного набора проводили в 10 повторениях, а для образцов из тестового набора в 5 повторениях.

При измерении ИК-спектров гречневой крупы (п. 4.2) образцы измельчали в лабораторной мельнице с последующим просеиванием через сито с диаметром отверстий, равным 0,43 мм, при этом измерение ИК-спектра каждого образца проводили в 10-17 повторениях.

Второй использованный спектроскопический метод – *UV-VIS-NIR*-спектроскопия. Для измерения *UV-VIS-NIR*-спектров применяли портативном спектрометре *ASD TerraSpec® 4 Hi-Res (TerraSpec, США)* в диапазоне от 350 до 2500 нм (рис. 5). Ключевым преимуществом данного спектрометра является возможность измерения без пробоподготовки, измерении пробы большой массы и объема, а также сочетание видимого (*VIS*) и ближнего инфракрасного (*NIR*), которое позволяет, с одной стороны, определить по *VIS*-диапазону цветовые характеристики и соответственно центры окраски, а с другой – идентифицировать по *NIR*-диапазону функциональные группы компонентов химического состава.



Рисунок 5 – портативный спектрометр *ASD TerraSpec® 4 Hi-Res*

Источник: составлено автором

UV-VIS-NIR-спектроскопия применялась при разработке моделей классификации гречневой крупы (п. 3.2). Использование портативного спектрометра для измерения *UV-VIS-NIR*-спектров образцов гречневой крупы имеет ряд преимуществ: измерение диффузного отражения можно проводить на месте («*in situ*») без пробоподготовки; при измерении спектров используется проба большого объема и массы, что повышает достоверность и репрезентативность результатов; сочетание видимого (*VIS*) и ближнего инфракрасного (*NIR*) диапазона позволяет, с одной стороны, определить по *VIS*-диапазону цветовые характеристики и соответственно центры окраски, а с другой – идентифицировать по *NIR*-диапазону функциональные группы компонентов химического состава гречневой крупы. Измерение *UV-VIS-NIR*-спектров для образцов гречневой крупы проводилось в 2-5 повторениях.

Спектроколориметрия. Спектры диффузного отражения образцов гречневой крупы (п.3.1.) в диапазоне 380-750 нм с интервалом 10 нм, с геометрией измерения $d/8$

измеряли на спектроколориметре *Color i5 (X-Rite Incorporated, USA)*, снабженного интегрирующей сферой и встроенной ксеноновой лампой, воспроизводящей источник света *D65* (рис. 6). Цветовые координаты определяли с помощью встроенной программы *COLOR iQs* при источнике освещения *D65* для колориметрического наблюдателя *CIE 1964* равном 10 градусам. Образцы помещали в кювету с длиной оптического пути 20 мм. Рассчитаны цветовые координаты в колориметрической системе *CIEL*a*b** (L^* -ось светлота; ось a^* - $+a^*$ -красный, $-a^*$ - зеленый; ось b^* - $+b^*$ -желтый, $-b^*$ синий).



Рисунок 6 – Спектроколориметр *Color i5 X-Rite*

Источник: составлено автором

Гиперспектральное изображение. Гиперспектральные изображения получали с помощью гиперспектральной камеры *Specim IQ (Spectral Imaging Ltd, Финляндия)* в спектральном диапазоне 400 - 1000 нм, (204 полосы) (рис. 7А). Для получения изображений область съемки была подсвечена галогеновой лампой *QL-500BW*. Гиперспектральные изображения также получали с помощью гиперспектральной камеры *Specim FX17e (Spectral Imaging Ltd, Финляндия)* в диапазоне измерений 935-1720 нм, спектральное разрешение 3,5 нм (рис. 7Б).



Рисунок 7 – Гиперспектральные камеры: А) *Specim IQ*, Б) *Specim FX17e*

Источник: составлено автором по данным [21]

Особенность этих устройств заключается в регистрации спектральных характеристик сразу для нескольких объектов исследования в множестве точек-пикселей. После измерения получается 3-мерная матрица данных, где координаты x и y соответствуют пикселям на изображении, а координата z – спектральным сигнатурам в определенных пикселях. Графическое представление метода гиперспектрального изображения представлено на рис. 8.

Гиперспектральная камера *Specim FX17e* была использована при разработке моделей классификации зерна гречихи по степени выполненности (п. 5.1). Из полученного гиперспектрального изображения отбирали спектральные сигнатуры, которые в дальнейшем использовали для разработки моделей классификации.

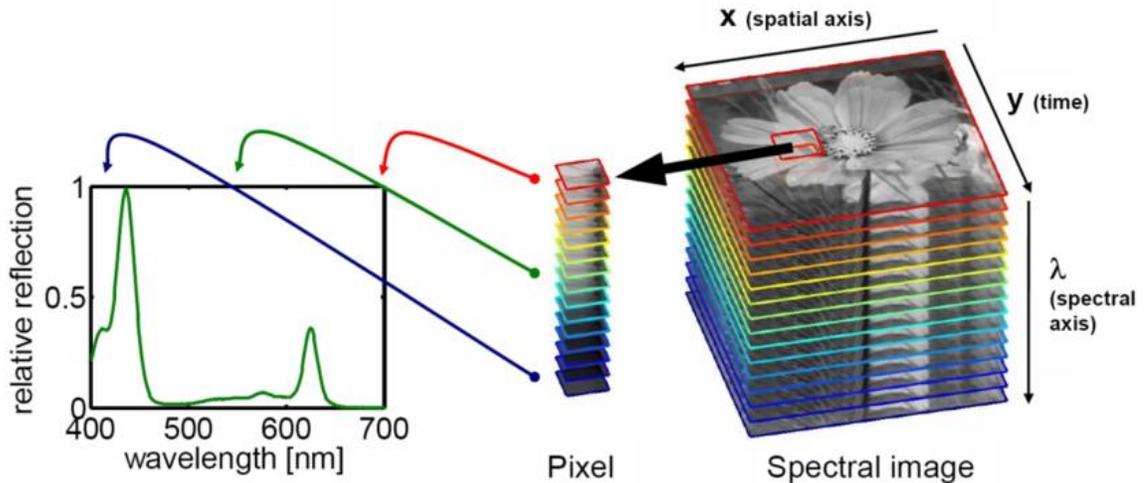


Рисунок 8 – Графическое представление гиперспектрального куба данных

Источник: составлено автором по данным [165]

При исследовании плодов авокадо (п. 5.2-5.4) регистрировали гиперспектральные изображения с использованием камеры *Specim IQ*. Для равномерного освещения образцов при регистрации изображений область съемки, содержащую образцы авокадо на листе коричневого картона, подсвечивали галогеновой лампой *QL 500BW Falcon Eyes* под углом освещения 30° . Гиперспектральная камера была установлена на штатив и размещена в верхней части освещаемой лампой области на расстоянии 30 см от образцов авокадо. При регистрации *HSI* с использованием *Specim IQ* происходит направление гиперспектральной камеры с помощью камеры-видеоискателя (*RGB*-камера). Аналогичное построение базы данных гиперспектральных изображений приведено в исследовании [173]. Полученные гиперспектральные изображения были импортированы в компьютер для последующей обработки.

Вместе с тем, при использовании гиперспектральных изображений отработан методический подход обработки и анализа изображений, который включает сегментацию гиперспектральных изображений с использованием выделения *ROI* (весь плод,

эллипсы в виде окружностей вокруг места дефекта) и спектральных сигнатур, предварительную обработку спектральных данных. Сигнатурой пикселя называют соответствующий ему вектор значений пикселя по спектральному диапазону.

Методы контрольного анализа. Определение кислотного числа жира (КЧЖ) образцов гречневой крупы проводили по стандартной методике (ГОСТ 31700-2012).

Измерение влажности ($MS, \%$) образцов плодов авокадо проводили с использованием влагомера *Mettler Toledo HB43-S* (рис. 9).



Рисунок 9 – Влагомер *Mettler Toledo HB43-S*

Источник: составлено автором

Для каждого образца авокадо из калибровочного набора отбирали пробу массой 10 г и высушивали при 100°C до постоянной массы. Разница между первоначальной массой W_0 и массой после высушивания W_1 использовалась для расчета сухого вещества ($DM, \%$) плодов авокадо из калибровочного набора, как представлено в уравнении (2).

$$DM(\%) = \frac{W_0 - W_1}{W_0} \times 100, \quad (2)$$

где W_0 – первоначальная масса;

W_1 – масса после высушивания.

При исследовании зерна гречихи была также применен рентгенографический анализ. Использована передвижная рентгенодиагностическая установка ПРДУ-02 (ЗАО «ЭЛТЕХ-Мед», Россия). Снимок получен с помощью считывателя рентгеновского изображения *DIGORA* (*Soredex*, Финляндия) в лаборатории технологии длительного хранения продовольственных товаров и хлебопродуктов ФГБУ НИИПХ Росрезерва [3].

2.3.2 Методы многомерного анализа

В диссертационной работе применялись методы многомерного анализа, направленные на решение следующих задач:

- 1) исследование данных с целью классификации и дискриминации;
- 2) прогнозирование новых значений.

Решение задач первой группы проводилось при наличии одного блока данных (спектры, гиперспектральные изображения), второй группы – при наличии двух блоков данных: предикторов и откликов [26].

В решении задач первой группы были задействованы многомерные методы: МГК, КА, ДА, *PCA-LDA*, *PCA-QDA*, *PLS-DA*, *SIMCA*.

Метод главных компонент применялся в данной работе для поиска критериев идентификации различных образцов пищевой продукции. При декомпозиции исходной матрицы на матрицы счетов и нагрузок происходит отбор наиболее статистически значимых переменных – отсчетов на частотах спектра, а также объяснение влияния максимальных значений нагрузок на распределение образцов на графике счетов (модель «без обучения»). Отобранные таким образом пики поглощения далее подвергаются проверке с использованием дискриминантного анализа (модель «с обучением»).

При использовании многомерных методов анализа для оценки репрезентативности используемых данных использовали сочетание параметров статистика Хотеллинга T^2 и F -статистики. Сочетание таких параметров позволяет оценивать используемые при разработке компьютерных моделей данные с процедурой удаления «выбросов».

В решении задач прогнозирования свойств пищевых товаров применялись следующие методы: РГК, ПЛС, SVR. Проверка качества и точности калибровочных моделей осуществлялась методом полной перекрестной валидации с использованием статистических показателей: R^2 (коэффициент детерминации), $RMSEC$, $RMSEV$ (*Root Mean Square Error*, среднеквадратичная ошибка калибровки и валидации соответственно) [25].

Применение методов многомерного анализа проводилось с использованием следующего программного обеспечения: *MS Excel*, *Unscrambler X 10.0.4.*, *Matlab 2018b*, *Prediktera Breeze ver. 2021.1.*, *Origin 2018b*, *Statistica ver. 10*.

Методы предварительной обработки спектральных данных. При применении многомерных методов также использовали методы предварительной обработки спектральных данных. Такие методы позволяют избавиться от фоновых шумов, возникающих при регистрации спектров, увеличить соотношение сигнал/шум, а также выявить скрытые различия в спектрах образцов. Использованные методы предварительной обработки спектральных данных включали следующее: выбор оптимального, репрезентативного диапазона спектра; усреднение спектральных данных; преобразование спектров отражения в спектры поглощения по формуле $A = \log I/R$; сглаживание спектральных данных методами скользящего среднего, Савицкого-Голея и др.; MSC ; SNV -коррекция (*Standard Normal Variate*).

Блок-схема основных этапов разработки моделей идентификации хемометрическими методами многомерного анализа представлена на рисунке 10.

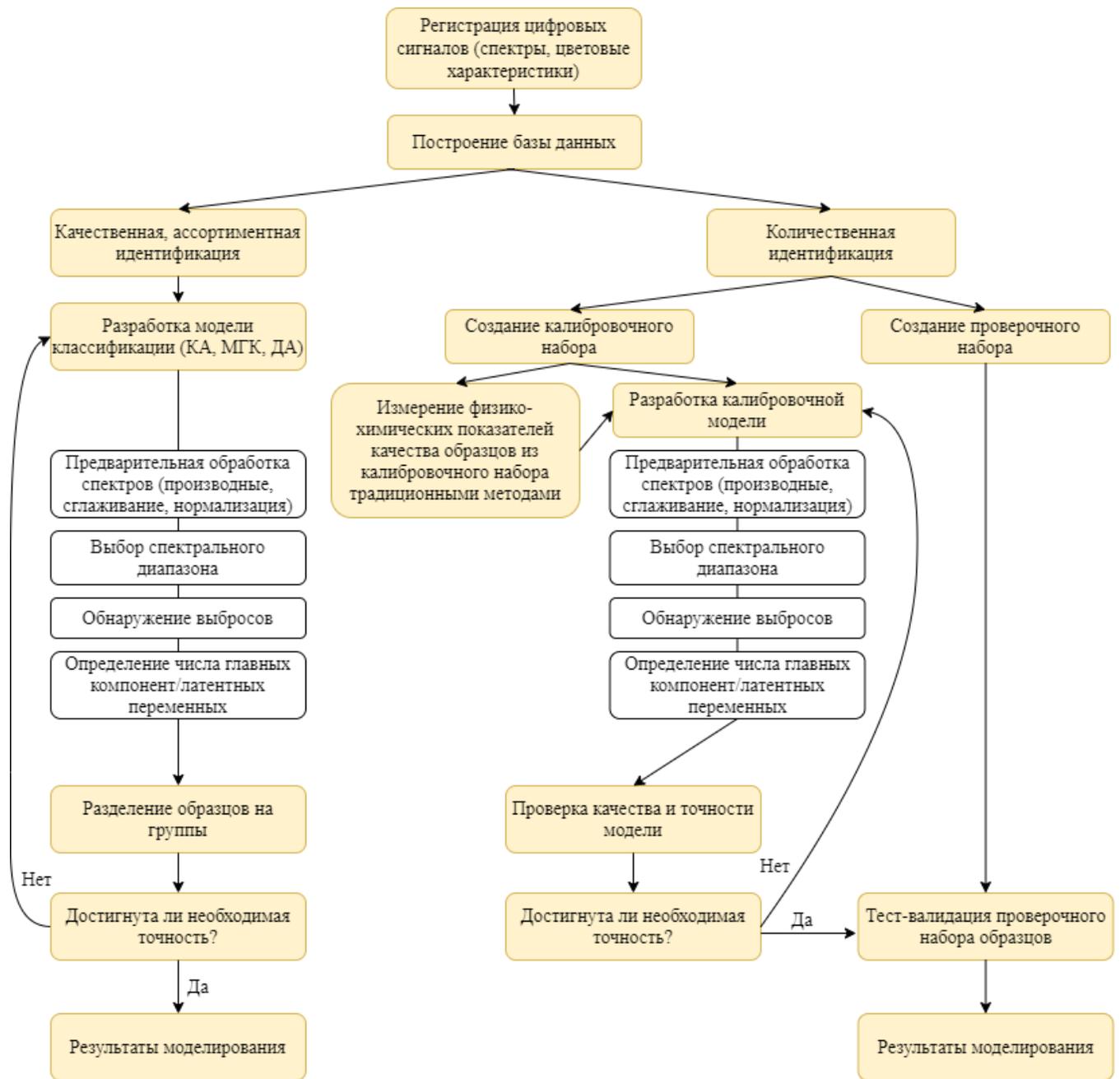


Рисунок 10 – Этапы хемометрического моделирования спектроскопических данных для классификации и прогнозирования свойств пищевых товаров

Источник: составлено автором

Глава 3 Разработка моделей классификации пищевых продуктов на категории качества

3.1 Градация гречневой крупы по данным спектроколориметрии методом дискриминантного анализа

Для проведения исследования были использованы образцы гречневой крупы, описанные в табл. 3. Для каждого образца измерены координаты цвета в системе $CIEL^*a^*b^*$ с помощью спектроколориметра *Color i5* (*X-Rite Incorporated*, США). Применение многомерных методов классификации осуществлялось в пакете программ *STATISTICA10* [24].

Градация гречневой крупы по цветовым координатам. Процедура «без обучения» - кластер-анализа и нахождение групп схожих объектов в выборке данных по множеству признаков. Расчеты были проведены методом агломеративного иерархического КА с использованием правила объединения между объектами – «методом полных связей», а в качестве меры сходства – евклидово расстояние между объектами (рис. 11). Для расчета была использована матрица $M \times N$, где M – цветовые координаты $L^*a^*b^*$ ($M=3$), N – образцы пропаренной и непропаренной гречневой крупы, отличающейся по срокам хранения ($n=20$).

По результатам кластерного анализа и при сопоставлении объектов кластеров с их значениями цветовых координат, согласно правилу Уорда определено число кластеров (рис. 11), образцы разделяются на четыре кластера (далее группы) по способу обработки (пропаренная и непропаренная) и сроку хранения [24].

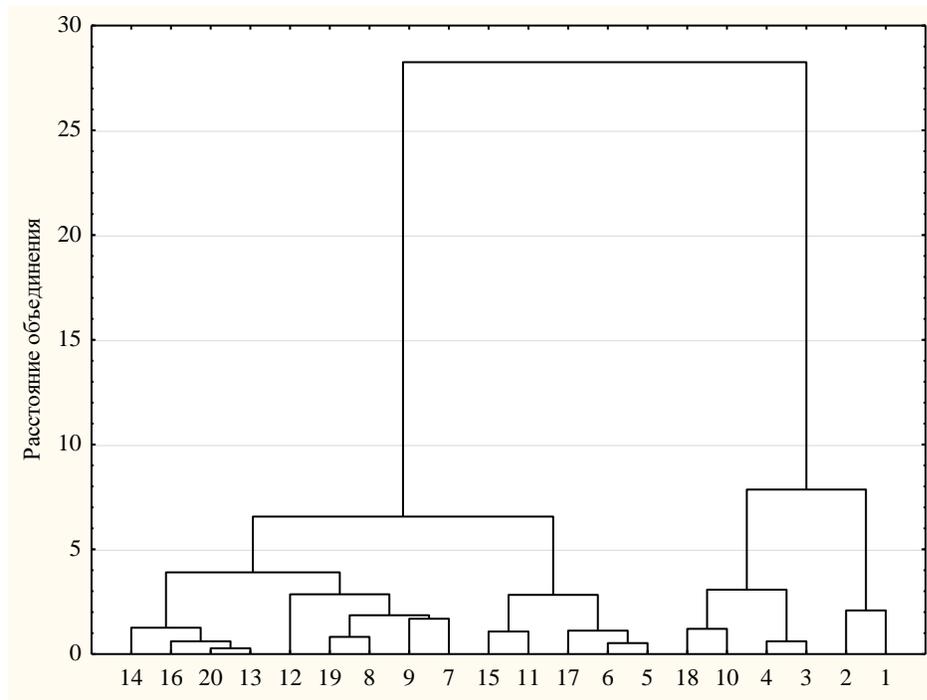


Рисунок 11 – Дендрограмма метода полных связей, с использованием евклидова расстояния, спектроколориметрической градации образцов гречневой крупы (кластер I - непропаренная гречневая крупа после года хранения; кластер II – непропаренная крупа; III и IV – пропаренная)

Источник: составлено автором

Методом главных компонент матрица $X=K \times N$, где K – коэффициент спектра отражения; N – число образцов крупы, разлагается в произведение двух матриц: T – матрица счетов и P – матрица факторных нагрузок. Число столбцов в матрице T и P , равное A , называют числом главных компонент. Из матриц выделено две главные компоненты, которые описывают 99,7% от общей дисперсии.

По данным T -матрицы счетов проведена градация с учетом знака (+/-) и значения координат образцов по 1 и 2 ГК. Получены четыре группы аналогично данным кластерного анализа: по 1 ГК – образцы разделяются на непропаренную и пропаренную, а по 2 ГК – непропаренная до и после хранения, а пропаренная – на 2 группы [24].

Расположение факторных нагрузок, которые информируют о корреляции между коэффициентами спектра отражения и главными компонентами, согласно данным R-матрицы, в координатах 1 и 2 ГК представлено на рис. 12. Все факторные нагрузки со знаком (+) по 1 ГК разделены по 2 ГК в зависимости от знака (+/-) на два диапазона: от 380 до 620 нм с максимумом при 530 нм (зеленый) и от 620 до 720 нм с минимумом при 680 нм (красный). Кроме того, выявлено ряд очень небольших изгибов на линии факторных нагрузок коэффициентов спектра отражения со следующими максимумами: 410, 460 и 650 и 680 нм, которые соответствуют полосам поглощения рутина [10,24] и хлорофилла [55,87].

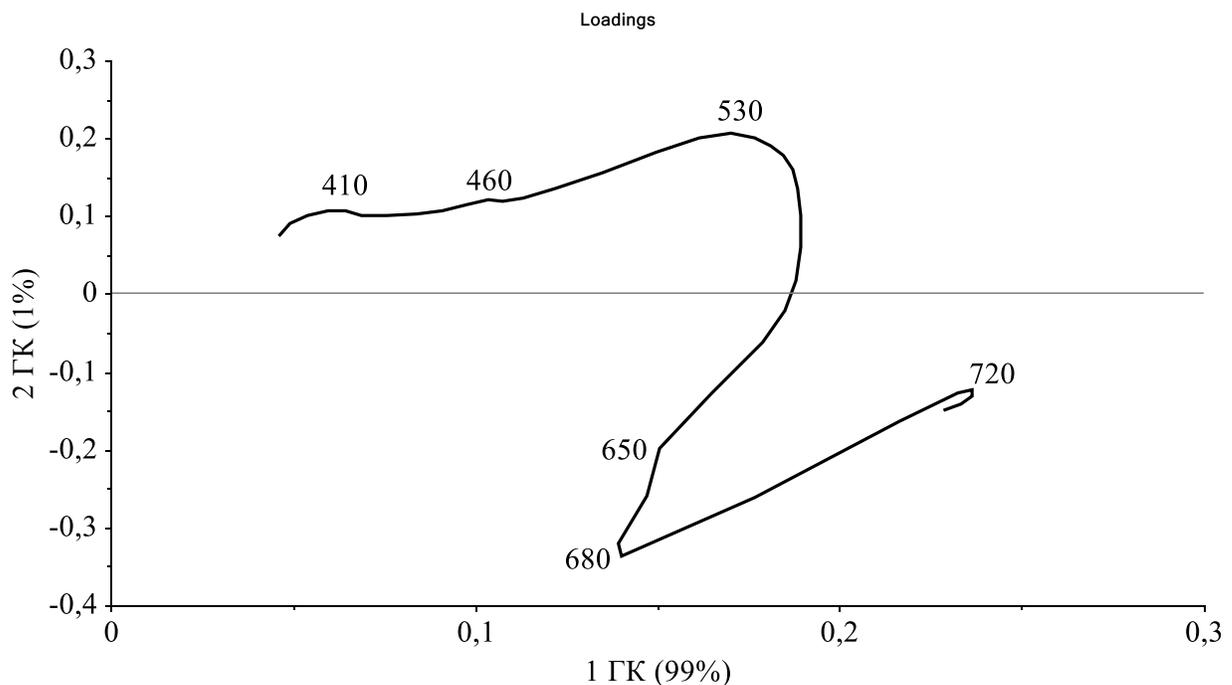


Рисунок 12 – Факторные нагрузки коэффициентов отражения спектра в координатах 1 и 2 ГК

Источник: составлено автором

Если сопоставить данные матриц T и P , то явно видно, что незначительные изменения значений и знака (+/-) по 1 и 2 ГК длин волн спектра, характерных для хлорофилла и рутина, свидетельствует об изменении их концентрации преимущественно после ГТО крупы.

Проверка качества многомерной модели классификации гречневой крупы по способу обработки (на пропаренную и непропаренную) и сроку хранения по цветовым координатам проводилась на обучающей выборке методом дискриминантного анализа.

Согласно статистики Уилкса (0,002) и F -статистики (9,34) = 44,84 (при уровне значимости $p < 0,0000$) различие по способу обработки (пропаренная и непропаренная) и сроку хранения значимо и наибольший вклад в цветоразличение гречневой крупы вносят координаты красноты a^* ($F=21,2$) и светлоты L^* ($F=13,2$) (таблица 4).

Таблица 4 – Статистическая проверка гипотезы о градации образцов гречневой крупы по цветовым координатам

Итоги анализа дискриминантных функций						
Показателей в модели: 3 Группы: 4						
Лямбда Уилкса: 0,002 $F(9,34)=44,84$ $p < 0,0000$						
Показатели	Лямбда Уилкса	Частная Лямбда	F включит. (3,14)	p -уровень	Толер.	1-толер. (R-кв.)
L^*	0,0078	0,26	13,2	0,000	0,68	0,322
a^*	0,0112	0,18	21,2	0,000	0,65	0,352
b^*	0,0023	0,87	0,68	0,579	0,91	0,087

Источник: составлено автором

В таблице 5 указаны средние значения цветовых координат образцов гречневой крупы, которые позволяют предварительно принять решение о принадлежности нового образца к той или иной группе [24].

Таблица 5 – Средние значения цветовых координат для центроидов образцов гречневой крупы

Номер группы	Цветовые координаты			Кол-во образцов
	L^*	a^*	b^*	
1	59,9	7,94	24,61	2
2	63,7	3,78	24,59	4
3	44,1	11,47	21,39	5
4	40,6	11,23	20,23	9
Всех групп	48,0	9,47	21,83	20

Источник: составлено автором

Подтверждено разделение по цветовым координатам $CIEL^*a^*b^*$ на четыре группы множества образцов гречневой крупы методом дискриминантного анализа с высоким уровнем значимости (табл. 4). Первая и вторая группы соответствуют координатам образцов непропаренной крупы, третья и четвертая – пропаренной (табл. 5).

В первой и второй группе образцы непропаренной гречневой крупы имеют светло зеленовато-желтый цвет, что обусловлено присутствием в нем рутина и хлорофилла, которые имеют полосы поглощения в сине-зеленом и красном диапазоне видимого спектра, соответственно. Из координат цветности a^* и b^* доминирует желтизна ($b^*>0$), с небольшой долей красного ($a^*>0$). В первой группе образцы непропаренной гречневой крупы, которые хранились в течение года наиболее окрашенные, чем во второй. Это может быть объяснено протеканием процесса ферментативного окисления органических соединений, входящих в состав гречневой крупы. Результатом ферментативного потемнения или покоричневения [49,160] является изменение цвета крупы: повышается краснота a^* (с 3,78 до 7,94) и понижается светлота L^* (с 63,7 до 59,9), при неизменной желтизне b^* . Насыщенность C^* данных образцов крупы в области цветового пространства $CIEL^*a^*b^*$ увеличивается (от 24 до 26), а цветовой тон h_{ab} смещается от желтого к красновато-желтому [24].

Третья и четвертая группы объединяют образцы пропаренной гречневой крупы, которые имеют коричневый цвет, приобретенный во время технологической обработки в результате неферментативного потемнения или покоричневения [49,160], что связано в основном с реакциями карамелизации сахаров и меланоидинообразования. При термической обработке (путем прожарки или пропаривания при высоких температурах с дополнительной просушкой) светло зеленовато-желтые ядра крупы приобретают коричневую окраску, что говорит о снижении концентрации рутина и хлорофилла и образовании продуктов карамелизации и меланоидиновых пигментов, которые имеют окраску от красно-коричневой до темно-коричневой. В результате ГТО у образцов крупы значения красноты a^* увеличиваются (от 3,78 до 11,47) и уменьшаются значения светлоты L^* (с 63,7 до 40,6) и желтизны b^* (с 24,61 до 20,23). Насыщенность C^* данных образцов крупы уменьшается до 22, а цветовой тон h_{ab} продолжает смещаться дальше к красно-желтой зоне цветового пространства $CIEL^*a^*b^*$.

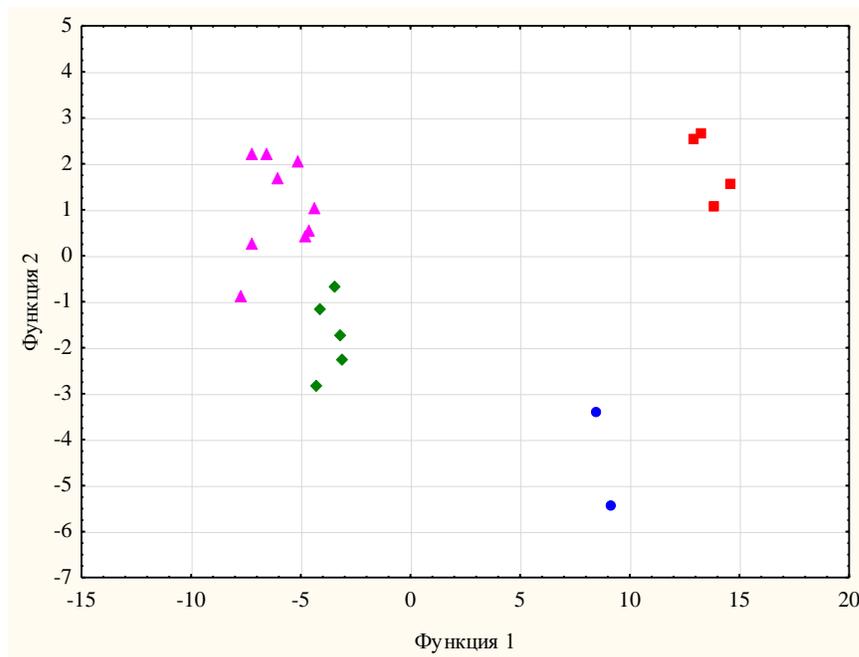


Рисунок 13 – Расположение точек, соответствующих образцам 1-4 групп гречневой крупы, в координатах двух дискриминантных функций (образцы гречневой крупы ○ – непропаренной после года хранения; □ – непропаренной; ◇ и △ – пропаренной)

Источник: составлено автором

Визуализация данных путем проектирования объектов – образцов гречневой крупы – в пространство двух дискриминантных осей отражена на рисунке 13. Множество образцов гречневой крупы из разных групп представлено как скопление точек в областях рассматриваемого пространства. Области четырех групп не пересекаются в этом пространстве, поскольку имеется значимое цветоразличение между ними.

Согласно знакам (+/+) и значений структурных коэффициентов координат цвета дискриминантной функции (ДФ), первая ДФ тесно связана с L^* - светлотой ($f = 0,91$) и b^* - желтизной ($f = +0,37$) против a^* - красноты ($f = -0,66$), а вторая ДФ с a^* - краснотой ($f = -0,70$). Следовательно, при повышении b^* - желтизны повышается светлота L^* , а при повышении a^* - красноты крупы резко снижается L^* (покоричневение крупы). Покоричневение крупы, как объяснено, связано с ферментативными и неферментативными процессами, а повышение желтизны – увеличение концентрации рутина [40] и хлорофилла в составе крупы.

Таким образом, полученные результаты свидетельствуют о перспективности применения спектроколориметрического метода и методов многомерного статистического анализа для быстрой и неразрушающей идентификации и контроля качества гречневой крупы.

3.2 Градация гречневой крупы по *UV-VIS-NIR*-спектрам и цветовым координатам

Гречневая крупа обладает значительной ролью в питании человека, что обусловлено содержанием сложных углеводов, белка, жиров, пищевых волокон и минеральных веществ [30,33,155,185]. Для увеличения объемов производства крупы при благоприятных климатических условиях возможен сбор урожая, перезимовавшего под снегом, весной. Несмотря на это, качество зерна, убранный весной, значительно

отличается от убранного осенью зерна [14]. Качество как зерна, так и крупы, убранного в осеннее или весеннее время, значительно различается по физическим и химическим показателям. В ранее проведенных исследованиях указывалось, что время сбора урожая влияет на количество проросших и испорченных зерен, влажность, кислотность и кислотное число жира [14,23].

Проблемы идентификации гречневой крупы и зерна исследовались в работах [13,15]. Изучено распределение ядер гречихи по фракциям и показано, что каждая фракция по крупности не может характеризоваться однородным составом [15], механические свойства перезимовавшего зерна изменяются неравномерно, соответственно такие данные необходимо принимать в расчет при шелушении и предпродажной подготовке зерна [13,23].

Вопросы идентификации гречневой крупы в зависимости от времени урожая и по размеру ядра в настоящее время изучены недостаточно. Для обеспечения эффективной сепарации, шелушения и дальнейшей обработки гречневой крупы, используемой в пищевых целях, необходимо: во-первых, обосновать взаимосвязь физико-химических свойств крупы со спектральными характеристиками, во-вторых, выявить полосы поглощения *UV-VIS-NIR*-спектра, вносящие наибольший вклад в разделение гречневой крупы по категориям качества [23].

Для исследований были использованы образцы гречневой крупы, описанные в таблице 3. Образцы пропаренной гречневой крупы осеннего и весеннего урожая из второго набора представлены в 6 фракциях по размеру ядра (производство ООО «Бийский элеватор», Алтайский край) [23] (табл. 6).

Таблица 6 – Различия образцов гречневой крупы различных фракций осеннего, весеннего урожая по размерам ядра и цветовым характеристикам

О – осенний урожай, V – весенний урожай

Условное обозначение	Размер ядра фракции гречневой крупы, мм	Цветовые координаты		
		L^* (O/V)	a^* (O/V)	b^* (O/V)
O6,V6	4,5	48,0/44,3	8,0/9,5	20,4/19,7
O5,V5	4,2	45,5/44,1	8,5/10,4	19,3/20,7
O4,V4	4,0	41,2/43,2	11,2/10,2	19,9/19,4
O3,V3	3,8	36,6/42,1	13,5/10,7	19,9/19,4
O2,V2	3,6	37,4/40,5	13,9/11,5	20,8/19,9
O1,V1	3,4	38,8/40,4	13,1/12,0	20,4/19,7

Источник: составлено автором

На основе измерений *UV-VIS-NIR*-спектров образцов была построена матрица данных X размерностью 130×2150 (130 – количество измерений спектров образцов, 2150 – отсчеты на длинах волн). Преобразование исходных *UV-VIS-NIR*-спектра образцов применялось для повышения качества и точности моделей. По *VIS*-диапазону спектра образцов рассчитаны цветовые характеристики в колориметрической системе *CIEL*a*b** [100].

Многомерные методы: иерархический агломеративный КА: правило объединения между объектами – «метод полных связей», мера сходства – евклидово расстояние; МГК: алгоритм – *NIPALS*, метод проверки – полная перекрестная валидация, применялись на основе программного комплекса *Unscrambler X 10.0.4 (Camo Software)*; ДА на основе программного комплекса *Origin 2019b (OriginLab Corporation)* [23].

До проведения многомерного моделирования проведена предварительная обработка *UV-VIS-NIR*-спектров: преобразование коэффициентов диффузного отражения (R) в коэффициенты длин волн спектра поглощения по функции $A = \log(1/R)$; *MSC*;

удаление частей спектра (350-380 и 2400-2500 нм) для уменьшения влияния шума; усреднение спектральных данных. Соответствие длин волн полос поглощения *UV-VIS-NIR*-спектра, приписываемых функциональным группам компонент и характеризующих химический состав образцов гречневой крупы, проведено по ранее опубликованным данным [23,55,64,80,94,107].

При анализе *UV-VIS-NIR*-спектров образцов гречневой крупы выявлено, что интенсивность полос поглощения при 1933 нм выше у образцов весеннего урожая по сравнению с образцами осеннего урожая, что свидетельствует о большем содержании в таких образцах влаги (рис. 14). В свою очередь, образцы осеннего урожая имели более высокую интенсивность поглощения при 1205, 1570, 1770, 2110 и 2310 нм, ответственным за содержание белка, крахмала и липидов. Интенсивность полосы с максимумом при 1460 нм изменяется от размера ядра и времени урожая крупы: при размере ядра крупы равном 4,2-4,5 мм ее интенсивность в среднем выше, чем при размере ядра ниже 4,2 мм; для образцов с размером ядра крупы 4,2-4,5 мм выявлено, что интенсивность этой полосы выше у образцов осеннего урожая, чем у образцов весеннего урожая [23].

Интенсивность полосы при 1933 нм (вода) выше для образцов непропаренной гречневой крупы, что связано с частичным разрушением ОН связей в процессе ГТО. Интенсивность полосы поглощения при 1205 нм (липиды, белок) у образцов непропаренной гречневой крупы также больше, что согласуется с [63]. Широкая и ассиметричная полоса при 1205 нм, приписываемая функциональной группе C=O, может соответствовать как липидам, так и белкам [94]. Среди образцов непропаренной гречневой крупы значения полос поглощения при 2310 и 2350 нм (липиды) меньше у образцов непропаренной гречневой крупы длительного хранения, что свидетельствует о протекании процесса окисления липидов при хранении [23]. Выявленные различия в интенсивности поглощения отдельных полос спектра приписываемым функциональным группам компонент состава гречневой крупы, соответствуют ранее опубликованным данным [14,23].

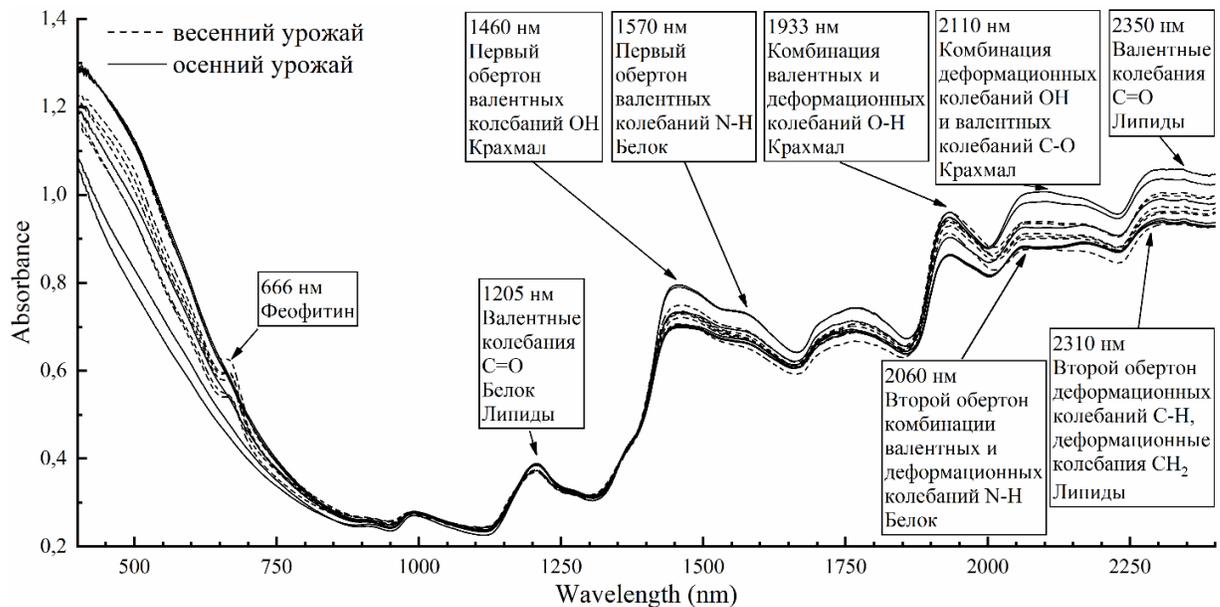


Рисунок 14 – Профиль спектров в *UV-VIS-NIR*-диапазоне (380-2400 нм) образцов весеннего и осеннего сбора

Источник: составлено автором

Поглощение в диапазоне длин волн от 350 до 840 нм спектра образцов непропаренной гречневой крупы более низкое по сравнению со спектрами образцов пропаренной гречневой крупой, что объясняет их более светлую окраску зеленовато-желтого тона. Изменение качественного состава хлорофилла гречневой крупы от времени урожая (осеннего и весеннего) и ГТО подтверждается разной формой спектра поглощения в диапазоне 640 – 680 нм (рис. 15). В спектрах непропаренной гречневой крупы проявляются две полосы поглощения с максимумами при 650 и 678 нм, что соответствует хлорофиллу *b* и *a*, соответственно (рис. 15в). В образцах пропаренной крупы из коммерческого набора (с номерами 5 – 17, 19 – 20) и весеннего урожая (с номерами VI – Vб) присутствует полоса поглощения феофитина *a* при 666 нм [124], образовавшегося из хлорофилла *a* при пропаривании, и отмечается потеря вспомогательного пигмента хлорофилла *b* [28,102] (рис. 15а). В спектрах образцов пропаренной крупы осеннего урожая полоса поглощения в этом диапазоне отсутствует [23] (рис. 15б).

Различия по цветовым характеристикам между образцами пропаренной и непропаренной гречневой крупы, включая непропаренную гречневую крупу с длительным сроком хранения, проанализированы ранее [23]. Координаты цвета в системе *CIE L*a*b* образцов гречневой крупы различных фракций весеннего и осеннего урожая приведены в табл. 6. По координате желтизна b^* образцы имеют незначительное различие от времени урожая и фракции крупы. С увеличением размера крупы отмечается повышение светлоты L^* : для образцов осеннего урожая от 38,8 до 48,0, а для образцов весеннего урожая - от 40,4 до 44,3 [23].

Наибольший вклад в цветоразличие пропаренной гречневой крупы разных фракций вносит координата краснота a^* , размах значений которой больше для образцов осеннего урожая ($\Delta a^* = 5,1$) в отличие от образцов весеннего урожая ($\Delta a^* = 2,5$). Коричневая окраска пропаренной гречневой крупы, фиксируемая через повышение значений координат цветности, преимущественно красноты ($+a^*$), связана, как с реакциями карамелизации сахаров и меланоидинообразования [23], так, вероятно, и с образованием из хлорофилла *a* феофитина *a* [23].

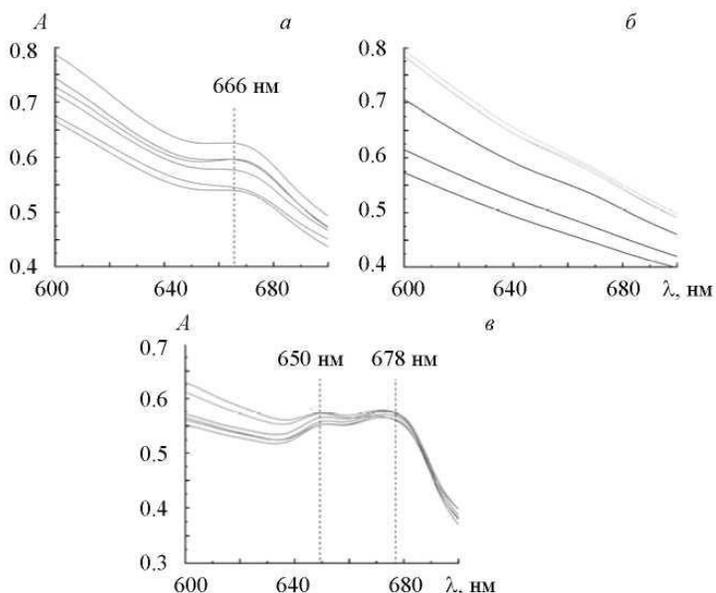


Рисунок 15 – *UV-VIS-NIR*-спектры образцов гречневой крупы в диапазоне 640 – 680 нм: весеннего урожая (а), осеннего урожая (б), непропаренной гречневой крупы (в)
Источник: составлено автором

По матрице X усредненных $UV-VIS-NIR$ -спектров методом КА выявлено 7 групп в зависимости от ГТО, срока хранения непропаренной гречневой крупы, размера ядра и времени урожая: группа 1 – непропаренная гречневая крупа (3,4,10,18); группа 2 – непропаренная гречневая крупа длительного хранения (1,2); группа 3 – гречневая крупа осеннего сбора размером ядра 4,2-4,5 мм (O_6, O_5); группа 4 – гречневая крупа весеннего сбора размером ядра 4,2-4,5 мм (V_6, V_5) и образцы пропаренной гречневой крупы (5,6,20); группа 5 – образец осеннего сбора размером ядра 4,0 мм (O_4) и образцы пропаренной гречневой крупы (8,9,11,13,14,17,19); группа 6 – образцы весеннего сбора размером ядра 3,6-4,0 мм (V_2, V_3, V_4) и образцы пропаренной гречневой крупы (15,16); группа 7 – образец весеннего сбора размером ядра 3,4 мм (V_1), образцы осеннего сбора размером ядра 3,4-3,8 мм (O_1, O_2, O_3) и образцы пропаренной гречневой крупы (7,12) [23].

Проведена декомпозиция X матрицы усредненных $UV-VIS-NIR$ -спектров (после их предварительной обработки) образцов гречневой крупы ($n=32$) на T матрицу счетов – координат образцов по 1 и 2 ГК (рис.16а), и P матрицу – факторных нагрузок длин волн спектра по 1 и 2 ГК (рис. 16б). Две главные компоненты объясняют 95% от общей дисперсии (87 и 8% соответственно). Расположение координат образцов и значения факторных нагрузок длин волн $UV-VIS-NIR$ -спектра поглощения в плоскости 1 и 2 ГК представлено на рис.16а и 16б. По максимальным значениям факторных нагрузок длин волн полос поглощения $UV-VIS-NIR$ -спектра, которые информируют об их корреляции с ГК, определен их вклад в разделение множества образцов на группы [23] (табл. 7).

Интерпретация двух ГК проведена по знаку (+/-) и значениям факторных нагрузок длин волн спектра. Согласно знаку (+/-) и значениям факторных нагрузок длин волн по 1ГК спектр разделен на 2 диапазона: знак (+) – имеют факторные нагрузки длин волн спектра, соответствующие видимому (VIS) диапазону, а знак (-) – ближнему инфракрасному (NIR) диапазону. Предложена следующая интерпретация главных компонент: 1 ГК объясняет степень зрелости пропаренной крупы от размера ядра, а 2

ГК - разделяет образцы от времени урожая (осеннего и весеннего) и ГТО (непропаренная и пропаренная) [23].

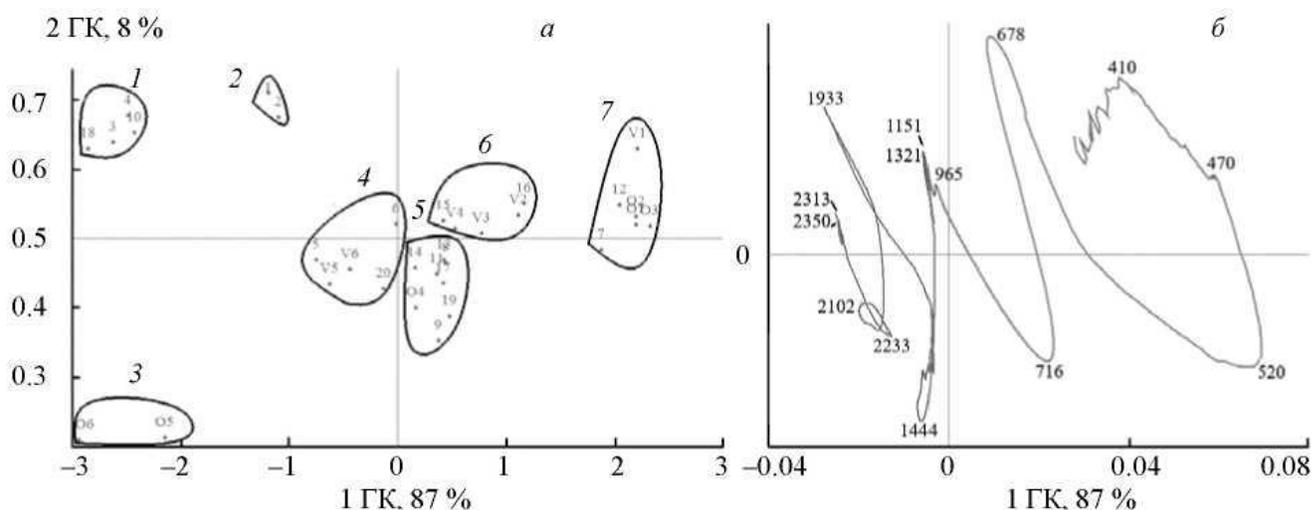


Рисунок 16 – Расположение точек, соответствующих образцам гречневой крупы, в координатах 1 и 2 ГК (а); факторные нагрузки длин волн *UV-VIS-NIR*-спектра поглощения в координатах 1 и 2 ГК (б)

Источник: составлено автором

Факторные нагрузки длин волн со знаком (+), в виде широкой полосы спектра по 1 ГК, имеет два максимума, различающиеся по знаку (-/+ по 2 ГК: первый – при 520, второй – при 678 нм, которые приписывают двум центрам окраски. Первый из центров окраски, расположенный в *UV*-диапазоне – представлен широкой полосой с максимумом факторной нагрузки при 520 нм, которая захватывает видимую область, вплоть до 716 нм (рис. 16б). Эта полоса поглощения (табл. 7) связана в основном с реакциями карамелизации сахаров и меланоидинообразования в процессе ГТО [23]. Второй центр окраски с максимумом значения факторной нагрузки длины волны при 678 нм соответствует полосе поглощения хлорофилла [23].

Ряд полос в *NIR*-диапазоне спектра со знаком (-) по 1ГК, которые различаются по знаку (+/-) по 2ГК, приписывают функциональным группам компонент, характеризующих химический состав гречневой крупы [23] (табл. 7).

Таблица 7 – Факторные нагрузки длин волн *UV-VIS-NIR*-спектра поглощения модели МГК

Максимум, нм	1ГК	2ГК	Функциональная группа	Компонент	Источник
Группы 1, 2					
678	0,0091	0,0567	-	Хлорофилл <i>a</i>	[55, 80]
965	-0,0029	0,0181	Второй обертон валентных колебаний ОН (абсорбционная вода)	Вода	[94]
1151	-0,0055	0,0270	C=O	Липиды	[64]
1321	-0,0055	0,0263	Колебания метальных групп CH ₃	Липиды	[64]
1933	-0,0274	0,0383	Комбинация валентных колебаний ОН и деформационных колебаний НОН	Вода	[94]
2313	-0,0243	0,0079	Второй обертон деформационных колебаний C-H, деформационные колебания СНг	Липиды	[64]
2350	-0,0237	0,0057	Валентные колебания C=O	Липиды	[64]
Группы 3, 4					
1444	-0,0060	-0,0434	Первый обертон валентных колебаний ОН	Крахмал (гидрок-сильная группа)	[94]
1560	-0,0049	-0,0328	Первый обертон валентных колебаний N-H	Белок	[94]
2042	-0,0154	-0,0198	Комбинация валентных колебаний N-H и C=O	Белок	[94]
2102	-0,0197	-0,0148	Комбинация деформационных колебаний ОН и валентных колебаний C-O Третий обертон валентных колебаний C-O	Крахмал	[64,101]
2233	-0,0127	-0,0212	Второй обертон деформационных колебаний N-H, комбинация валентных колебаний C-H и C=O	Белок	[64]
Группы 5, 6, 7					
716	0,0227	-0,0273	-	Продукты меланоидинообразования и карамелизации	[23]
520	0,0694	-0,0259	-		
470	0,0590	0,0204	-	Каротиноиды	[55, 80]

Источник: составлено автором по данным [23]

На рис. 16а показано расположение кластеров (групп) образцов гречневой крупы, полученных методом КА, в координатах двух ГК. Согласно значениям факторных нагрузок длин волн, приписываемые функциональным группам компонент состава гречневой крупы, в диапазоне выявлены те, которые вносят наибольший вклад в разделении образцов на группы [23] (табл. 7).

При сопоставлении расположения по знаку (+/-) факторных нагрузок длин волн спектра и координат образцов из разных групп в плоскости 1 и 2ГК выявлен ряд закономерностей. Во-первых, в зависимости от знака: от (+) до (-) координат образцов по 1ГК, увеличивается размер ядра, что характерно для пропаренной гречневой крупы осеннего и весеннего урожая (от 3 до 7 группы образцов крупы). Согласно [13,15], размер ядра гречневой крупы связан со зрелостью зерна. Расположение образцов пропаренной гречневой крупы осеннего и весеннего урожая от $O1/V1$ до $O6/V6$ (табл. 6), различающиеся по размеру ядра (от 3,4 до 4,5 нм) аналогично их расположению в плоскости координат цвета: светлоты L^* и красноты a^* . С увеличением размера ядра повышается светлота L^* (от 38,8/40,4 до 48,0/44,3) гречневой крупы при уменьшении красноты a^* (от 13,1/12,0 до 8,0/9,6). Этот порядок расположения в координатах двух систем хорошо согласуются с интерпретацией значений факторных нагрузок длин волн в диапазоне от 380 до 716 нм спектра поглощения (рис.16б). Во-вторых, образцы гречневой крупы в координатах 1 и 2ГК значимо различаются в зависимости от ГТО: на непропаренную (1 и 2 группы) и пропаренную (3 - 7 группы). Образцы непропаренной гречневой крупы, включенные в 1 и 2 группы, находятся в плоскости 1 и 2ГК со знаком (-/+), соответственно, аналогично знакам (-/+) факторных нагрузок длин волн при 965 и 1933 нм (влага) и 2315 и 2350 нм (липиды), что соответствует данным о большом содержании влаги и липидов [64]. При этом образцы из 1 и 2 группы значимо различаются по значениям координат 1 и 2 ГК, что объясняется протеканием процесса ферментативного окисления органических соединений при длительном хранении крупы из 2 группы, как указано в [23].

Группа 3, объединяющая образцы осеннего сбора с размером ядра 4,2 – 4,5 мм, расположены в координатах 1 и 2 ГК аналогично по знакам факторных нагрузок длин волн полос поглощения при 1444, 2042, 2102 и 2233 нм, приписываемые функциональным группам крахмала и белков (табл. 7). Поэтому образцы этой группы пропаренной крупы, в отличие от других групп, содержат больше углеводных и белковых соединений, что согласуется с данными [14]. Группа 4, включающая образцы весеннего сбора с размером ядра 4,2 – 4,5 мм, также характеризуется высоким содержанием углеводов и белков, но меньшим чем у образцов группы 3. Группы 5-7 пропаренной гречневой крупы взаимосвязаны с факторными нагрузками полос поглощения в диапазоне 380 -716 нм, значения которых обусловлены реакциями меланоидинообразования и карамелизации при ГТО [23,24].

Решение задачи идентификации и сортировки гречневой крупы подразумевает использование небольшого количества длин волн из широкого диапазона *UV-VIS-NIR* спектра, вносящих наибольший вклад в разделение образцов по категориям качества. С помощью методов КА и МГК проведено разделение образцов на 7 групп и определены длины волн с максимальным значением факторных нагрузок: 520, 678, 1444 1933 нм, которые были использованы в качестве дискриминантных переменных [23].

Идентификация гречневой крупы по качеству сведена к нахождению системы классификационных функций, полученных на обучающей выборке методом дискриминантного анализа. Для решения этой задачи использовали два блока данных: первый блок *X* – матрица размерностью 130 x 4, в которой 130 строк – количество образцов; 4 столбца – коэффициенты длин волн спектра поглощения; второй блок *Y* – матрица из 130 строк значений номеров групп для образцов. С помощью классификационной модели, связывающей блоки *X* и *Y*, можно в дальнейшем предсказать номер группы по значениям длин волн спектра поглощения нового образца [23].

По собственному значению классификационных функций, которые связаны с их дискриминантной возможностью (в % от общей), установлено, что основное различие – по первой и второй функции и менее значимо – по третьей (72,8%, 23,6% и 2,6%

соответственно). Третья функция, кроме низкой дискриминантной возможности, имеет более низкие значения канонической корреляции ($R^2 = 0,65$) против соответствующих значений первой и второй функций (0,98 и 0,93), поэтому, как следствие, не очень тесно связана с группами образцов и не является, соответственно, информативной. Общее количество правильно классифицированных образцов в каждой группе равно 100%. Следовательно, согласно результатам моделирования методом дискриминантного анализа подтверждено разделение образцов гречневой крупы на 7 групп [23].

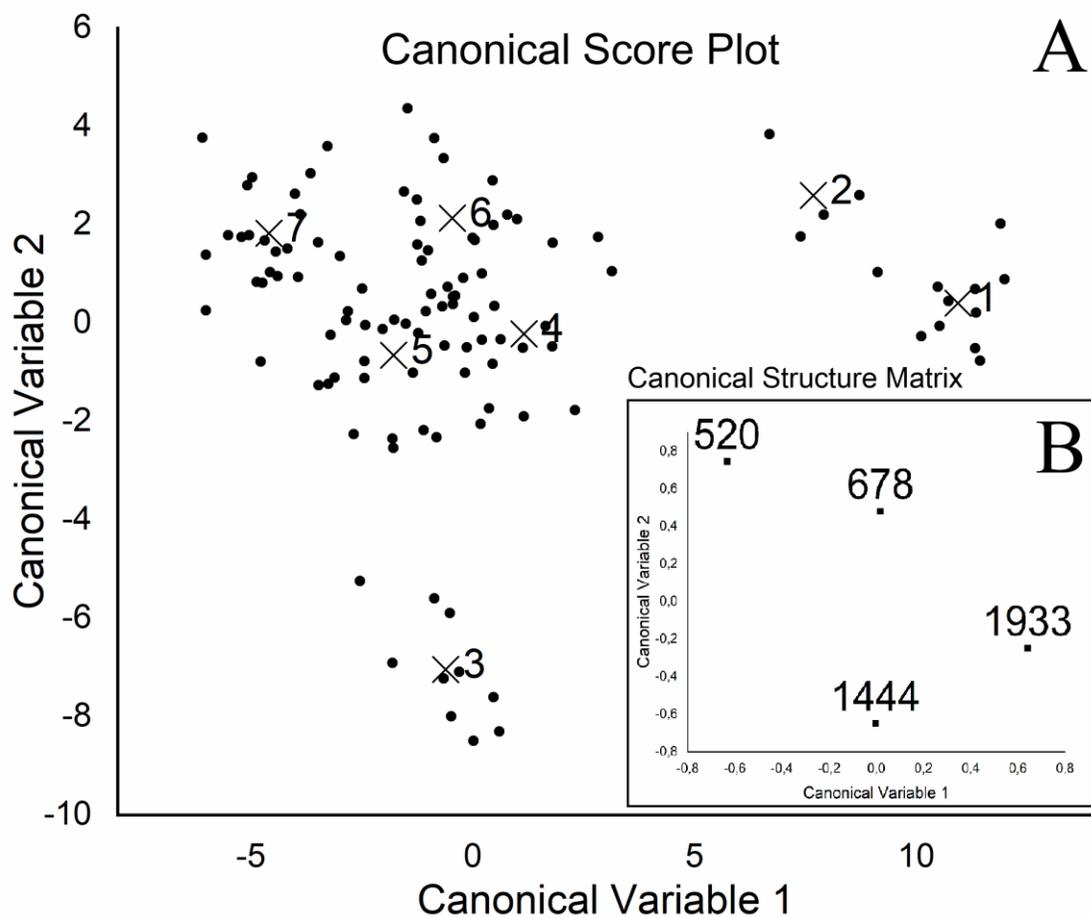


Рисунок 17 – Расположение точек, соответствующих образцам 1-7 групп гречневой крупы (А), и структурных коэффициентов (В) в координатах двух канонических дискриминантных переменных

Источник: составлено автором

Расположение координат объектов – образцов гречневой крупы - в пространстве двух канонических дискриминантных функций представлено на рис. 17А. Как видно из этого рисунка, области групп 1,2 и 3 имеют значимое различие в координатах двух функций, а более близкое расположение между группами 4-7 объясняется схожими спектральными характеристиками и, соответственно, химическим составом образцов пропаренной крупы. По первой функции, обладающей большей дискриминантной возможностью, образцы разделяются на группы непропаренной (1,2) и пропаренной (3-7) гречневой крупы, а по второй функции – на группу (3) пропаренной крупы с большим размером ядра осеннего урожая против групп (4-7) пропаренной крупы. Градация образцов на группы объясняется различием по знаку (+/-) и значению структурных коэффициентов (520, 678, 1444 и 1933) двух дискриминантных функций [23] (рис. 17В).

Методом канонического дискриминантного анализа используя два блока X и Y данных, рассчитаны константы b_{ko} и коэффициенты b_{kl}, \dots, b_{kp} , соответствующие классификационным функциям h_k для семи выделенных групп образцов гречневой крупы. Система классификационных функций представлена в виде:

$$h_k = b_{ko} + b_{k1}X_1 + \dots, b_{kp}X_p, \quad (3)$$

где h_k – значения классификационной функции для K-группы ($K = 1, \dots, 7$) образцов;

b_{ko} – константа функции для K-группы;

b_{kl}, \dots, b_{kp} – коэффициенты функций K-группы для p- коэффициентов длин волн ($p=1,2,3,4$);

$X_1 \dots X_p$ – коэффициенты длин волн (520; 678; 1444 и 1933 нм) спектров поглощения образцов.

Таким образом, классификация образцов гречневой крупы, различающейся по размеру ядра, времени урожая, сроку хранения, в зависимости от ГТО, сведена к нахождению системы классификационных функций h_k , полученной по двум блокам

данных X и Y . Идентификация образцов гречневой крупы осуществляется по их принадлежностям к одной из семи групп, которая устанавливается по наибольшему значению классификационной функции h_k [23].

3.3 Идентификация зернового кофе методами ИК-Фурье спектроскопии и многомерного анализа

Наиболее популярными ботаническими видами кофе, которые реализуются на мировых и отечественных рынках, являются Арабика (*Arabica Linney*) и Робуста (*Canephora Pierre*) [39,43]. Изучены различия зернового кофе в зависимости от ботанического вида по химическому составу и показано [22], что кофе Арабика содержит больше липидов, хлорогеновой кислоты и сахаров, а Робуста – больше кофеина, белков и воды. Изменения химического состава кофе при обжарке показаны в [174]: при обжарке кофе уменьшается содержание поли- и олигосахаридов, свободных аминокислот, хлорогеновой кислоты, кофеина при увеличении содержания липидов, белков, минералов и меланоидинов. Также отмечается [61], что при обжарке кофе происходит деградация хлорогеновой кислоты и почти полная деградация сахарозы [128].

Сочетание ИК-Фурье спектроскопии методами многомерного анализа используют как быстрый и неразрушающий инструмент для количественного и качественного контроля качества кофе [35,42,53,54,76,98,112,128,158]. Поэтому разработка метода идентификации зернового кофе по ботаническому виду, географическому месту произрастания и обжарке применением методов ИК-Фурье спектроскопии и многомерного анализа является актуальной [128].

Для исследования были использованы образцы кофе, указанные в таблице 3. Образцам был присвоен индекс, включающий: 1) номер образца по порядку; 2) букву «А», «Р» перед номером образца, обозначающий ботанический вид Арабика или Робуста, соответственно; 3) не обжаренным образцам зернового кофе была присвоена

буква «н». В контрольную выборку были включены образцы обжаренного и не обжаренного декофеинизированного кофе, а также кристаллический кофеин (содержание кофеина 98,5%) [128].

По результатам измерения спектральные данные были экспортированы из встроенного программного обеспечения *OPUS 7.3.5* в *Excel*, в котором сформирована матрица данных, размером $X=I \times J$, I – количество усредненных ИК-спектров образцов кофе ($n=39$), а J – коэффициенты поглощения на волновых числах ИК-спектра ($n=1748$).

Анализ матрицы X данных проводили методом главных компонент в программном пакете *Unscrambler X (ver. 10.0.4, Camo Software, Норвегия)*. Дискриминантный анализ (ДА) выполнен с использованием программного комплекса *Origin 2018b (OriginLab Corporation, США)*.

При визуальном анализе ИК-спектров выявлен ряд полос поглощения, которые приписывают, согласно [35,42,53,54,76,98,112,128], функциональным группам следующих компонентов состава зернового кофе: сахара, углеводы, хлорогеновая кислота, кофеин, липиды и вода (табл. 8).

Таблица 8 – Назначение полос ИК-спектра функциональным группам компонентов состава зернового кофе

Полоса, см ⁻¹	Функциональная группа	Компонент	Источник
1021	Валентные колебания С-О	Целлюлоза, углеводы	[35,53,128]
1011			
1027			
1041			
1151	Валентные колебания С-О	Полисахариды, целлюлоза	[42]
1237	2-й обертона СН в СН ₂ и СН группах		[42]
1455	1-й обертона валентных колебаний ОН	Вода, хлорогеновая кислота	[53,128]
1515			[42]

Полоса, см ⁻¹	Функциональная группа	Компонент	Источник
1552	Симметричные деформационные колебания NH ₃ , полоса амида 2	Аминокислоты, белок	
1600	Циклические амиды	Кофеин	[42]
1647			
1744	Валентные колебания C=O	Липиды, альдегиды, алифатические эфиры	[98,112]
2853	Валентные колебания CH ₃	Метильные группы в молекулах кофеина и липидах	[54,128]
2921	Валентные колебания CH ₂		
3008	Валентные колебания C-H дис-олефиновых групп	Липиды	[53]
3284	Валентные колебания OH	Вода	[53]

Источник: составлено автором

Как показано в [35,53,128], основное различие по ИК-спектрам между образцами кофе выявлено в диапазоне от 700 до 1700 см⁻¹. Для сравнения ИК-спектров в этом диапазоне были выбраны следующие образцы: кофеин (*caffeine*), зерновое кофе ботанического вида Арабика (*arabica*), а также не обжаренное (*decaf*) и обжаренное (*decaf roasted*) декофеинизированное кофе (рис. 18). Полоса поглощения с максимумом при 759 см⁻¹, относящаяся к области отпечатков пальцев ИК-спектра, проявляется в спектрах кофеина, кофе (Арабика) и отсутствует в спектрах декофеинизированного кофе. Широкая полоса поглощения с максимумом при 1027 см⁻¹, соответствующая валентным колебаниям C-O [35,53,128], проявляется в ИК-спектрах образцов кофе и в виде узкой полосы у кофеина. В ИК-спектре кофеина отсутствует полоса поглощения при 1155 см⁻¹ (валентные колебания C-O в полисахаридах, целлюлозе) в отличие от спектров кофе. Во всех четырех ИК-спектрах образцов проявляется полоса поглощения при 1455 см⁻¹, соответствующая 1-му обертону валентных колебаний OH в воде и хлорогеновой кислоте [53,128]. Присутствие данной полосы в спектре кофеина может быть связано с незначительным присутствием влаги в порошке кристаллического ко-

феина. Полоса поглощения при 1647 см^{-1} , которая согласно [76] приписана колебаниям циклических амидов в молекуле кофеина, проявляется в ИК-спектрах кофеина, кофе Арабика и обжаренном декофеинизированном кофе. Поэтому данная полоса, может соответствовать поглощению не только кофеина, но и белков и свободных аминокислот [174], интенсивность которых, вероятно, увеличивается при обжарке. Следовательно, соотношение интенсивностей вышеуказанных полос поглощения ИК-спектра можно использовать для построения калибровочных моделей содержания кофеина в составе зернового кофе [128].

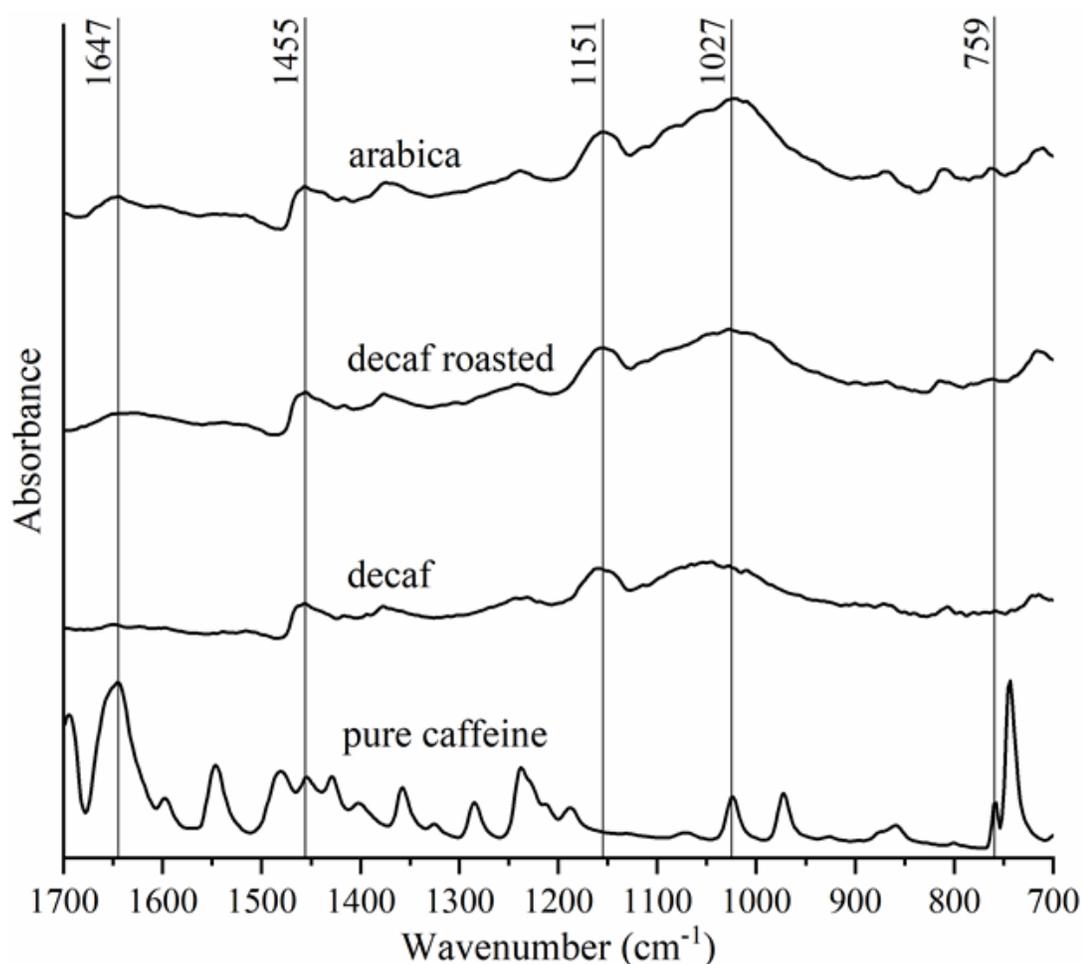


Рисунок 18 – Профили ИК-спектров образцов кофеина, обжаренного и не обжаренного декофеинизированного кофе и кофе вида арабика в диапазоне $1700\text{-}700\text{ см}^{-1}$

Источник: составлено автором

Важной проблемой применения методов многомерного анализа является решение задач качественной идентификации кофе, в которых требуется установить принадлежность образцов в зависимости от ботанического вида и обжарки.

Для построения многомерной модели классификации использовали матрицу X усредненных ИК-спектров образцов кофе. Декомпозиция матрицы X усредненных ИК-спектров образцов кофе на T -матрицу счетов и P -матрицу факторных нагрузок волновых чисел спектра проведена методом главных компонент. Три ГК объясняют 99% от общей дисперсии (84, 12 и 3% соответственно) [128].

По T -матрице счетов - координат образцов по каждой их трех ГК по знаку (+/-), выявлено разделение образцов кофе на группы (рис. 19). В координатах 1 и 2 ГК образцы кофе подразделяются от обжарки: не обжаренное и обжаренное (рис.19А), а в координатах 1 и 3 ГК – от ботанического вида (Арабика/Робуста) (рис. 19В). Факторные нагрузки, показывающие корреляцию коэффициентов поглощения на волновых числах спектра по 2 и 3 ГК, показаны на рис. 20.

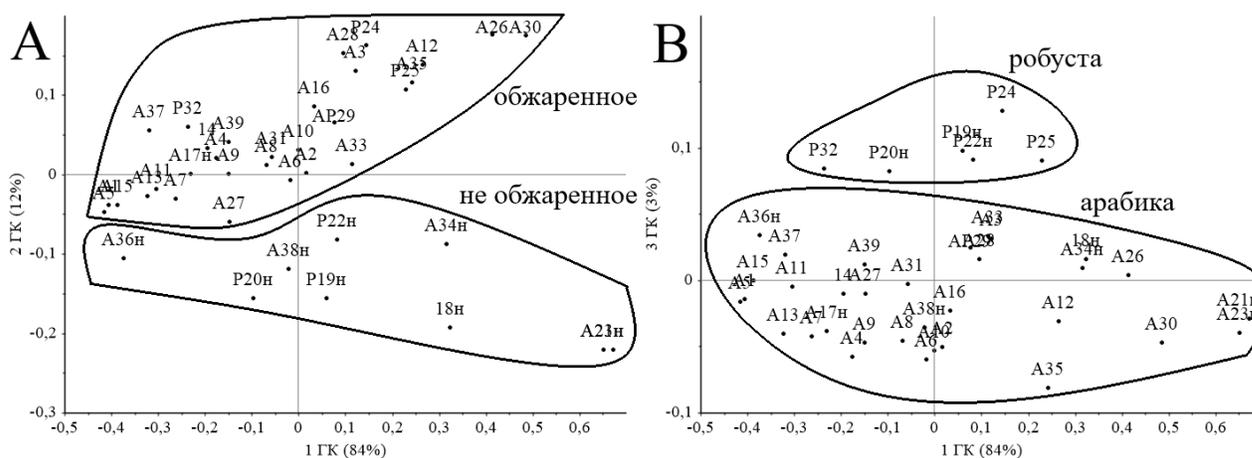


Рисунок 19 – Расположение точек, соответствующих образцам кофе, в координатах 1 ГК и 2 ГК (А), в координатах 1 ГК и 3 ГК (В)

Источник: составлено автором

Первая ГК, объясняющая 84% от общей дисперсии, тесно связана со всеми коэффициентами поглощения волновых чисел ИК-спектра, что подтверждает гипотезу

о слабом различии между образцами по химическому составу. Но при использовании соотношения значений факторных нагрузок коэффициентов поглощения на волновых числах спектра по 1 и 2 или 1 и 3 ГК выявлены те полосы ИК-спектра, которые вносят вклад в разделение образцов на группы [128] (рис. 20).

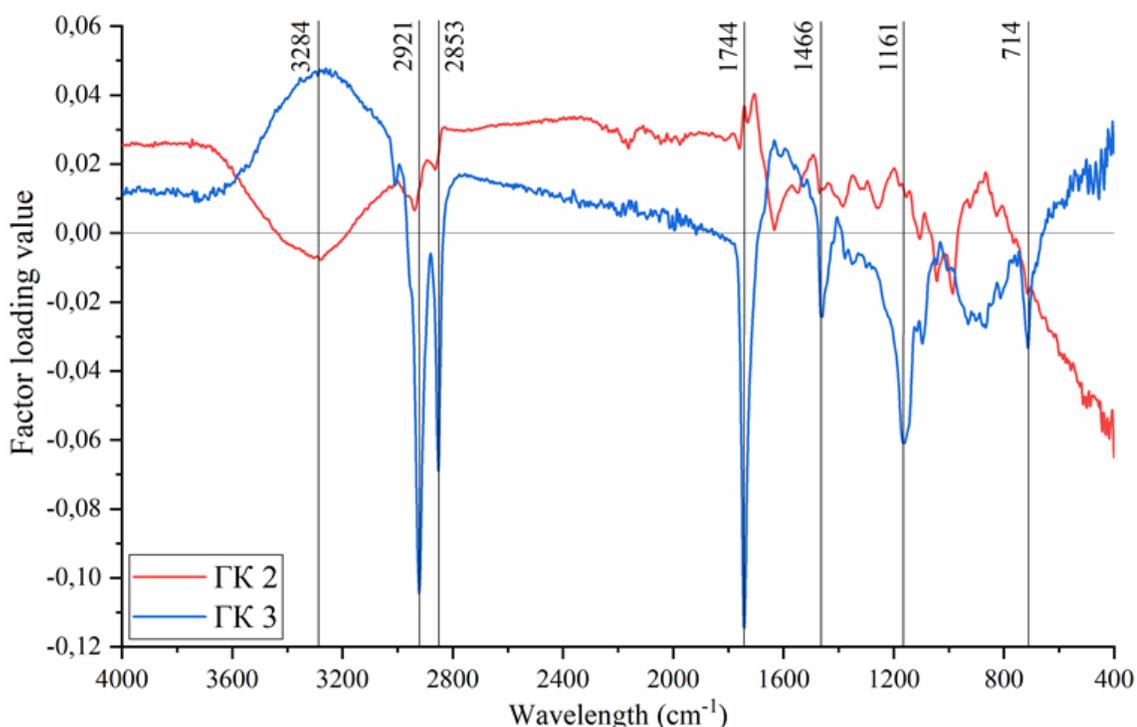


Рисунок 20 – Факторные нагрузки коэффициентов поглощения волновых чисел ИК-спектра по 2 и 3 ГК

Источник: составлено автором

В разделение образцов по 2 ГК: по обжарке, вносят полосы поглощения ИК-спектра, приписанные функциональным группам различных компонентов состава кофе. Максимальные значения факторных нагрузок со знаком (+) по 2 ГК имеет ряд полос: полоса при 1744 см^{-1} (липиды, альдегиды, алифатические эфиры) и полосы в диапазоне от 1500 до 1140 см^{-1} (полисахариды, хлорогеновая кислота). Различия в значениях факторных нагрузок связано с процессами, протекающими при обжарке кофе, а именно с увеличением липидов [174]. Ряд факторных нагрузок в диапазоне от 1500 до 1140 см^{-1} имеют знак (+), но направлены в область со знаком (-), что говорит о

процессах деградации хлорогеновой кислоты при обжарке [61]. Аналогично, факторные нагрузки в диапазоне от 1600 до 1647 см^{-1} (циклические амиды в молекуле кофеина) имеют знак (+), но направлены в область со знаком (-), соответствуют большему содержанию кофеина в не обжаренном кофе в отличие от обжаренного. Факторные нагрузки со знаком (-) в диапазоне от 1161 до 1000 см^{-1} (валентные колебания С-О) объясняют деградацию поли- и олигосахаридов в кофе при обжарке [174]. Высокое значение факторной нагрузки при 3284 см^{-1} (валентные колебания ОН) со знаком (-) объясняет уменьшение содержание воды при обжарке [61,128].

В разделении образцов по 3 ГК: по ботаническому виду, вносят полосы поглощения ИК-спектра, которые приписывают функциональным группам липидов, кофеина, хлорогеновой кислоты и углеводов. Значение факторной нагрузки при 3284 см^{-1} (валентные колебания ОН) со знаком (+) по 3 ГК соответствует большему содержанию воды в Робусте в отличие от Арабики. Факторные нагрузки со знаком (-) при 2921, 2853, 1744, 1466 и 1161 см^{-1} объясняют наличие большего количества липидов, хлорогеновой кислоты и углеводов в Арабике в отличие от Робусты [35,98,112,128]. Факторные нагрузки в диапазоне от 1600 до 1647 см^{-1} (циклические амиды в молекуле кофеина) имеют знак (+) и подтверждают наличие большего количества кофеина в Робусте в отличие от Арабики [76]. Факторная нагрузка при 1554 см^{-1} (симметричные деформационные колебания NH_3 молекул белка) со знаком (+) по 2 ГК показывает большее содержание белков в Робусте [42,61,128,174].

В последнее время определение географического места произрастания продуктов питания, включая зерновой кофе, имеет особое значение. Контроль аутентичности зернового кофе по географическому месту произрастания имеет определенную сложность, так как состав продукта и, следовательно, профиль ИК-спектров зернового кофе слабо зависит от его места произрастания. Для классификации кофе в зависимости от географического места произрастания в настоящее время используют сочетание ИК-спектроскопии и многомерного анализа.

Определение зернового кофе по географическому месту произрастания, в данном исследовании, сведена к нахождению системы классификационных функций, полученных на обучающей выборке методом дискриминантного анализа. В качестве переменных выбрано 19 коэффициентов поглощения на волновых числах спектра: 3284, 2921, 2853, 2851, 1744, 1742, 1704, 1645, 1634, 1597, 1459, 1163, 1155, 1095, 1041, 986, 864, 759, 712 см^{-1} , которые имели максимальные значениям факторных нагрузок в многомерной классификационной модели, полученной методом МГК [128].

Для построения модели идентификации кофе использовали два блока данных: первый блок матрица X размерностью 114×19 , где 114 строк – количество спектров образцов, а 19 столбцов – коэффициенты поглощения на волновых числах спектра; второй блок Y – матрица из 114 строк значений номера групп ($n=3$; 1 – Азия, 2 – Америка, 3 – Африка) образцов [128].

Согласно априорной информации о географическом месте произрастания определена дискриминантная возможность классификационных функций (в % от общей) и установлено, что основное различие происходит по первой дискриминантной функции (ДФ) и менее значимо по второй (66,4% и 33,6% соответственно). Общее количество правильно классифицированных образцов в каждой группе варьирует от 86,8 до 100% [128].

Интерпретация градации образцов на три группы проведена по знаку (+/–) и значениям структурных коэффициентов волновых чисел по 1 и 2 ДФ. Образцы из первой группы, согласно знаку (–) по 1 ДФ расположены против 2 и 3 групп, имеющих знак + по 1 ДФ, но они, в свою очередь, по знаку (+/–) разделяются по 2 ДФ (рис. 21).

Таким образом, определение географического места произрастания зернового кофе сведена к нахождению системы классификационных функций, полученной по двум блокам данных X и Y . Аутентификация образцов кофе проводится по их принадлежности к одной из трех групп по наибольшему значению по одной из трех функций [128].

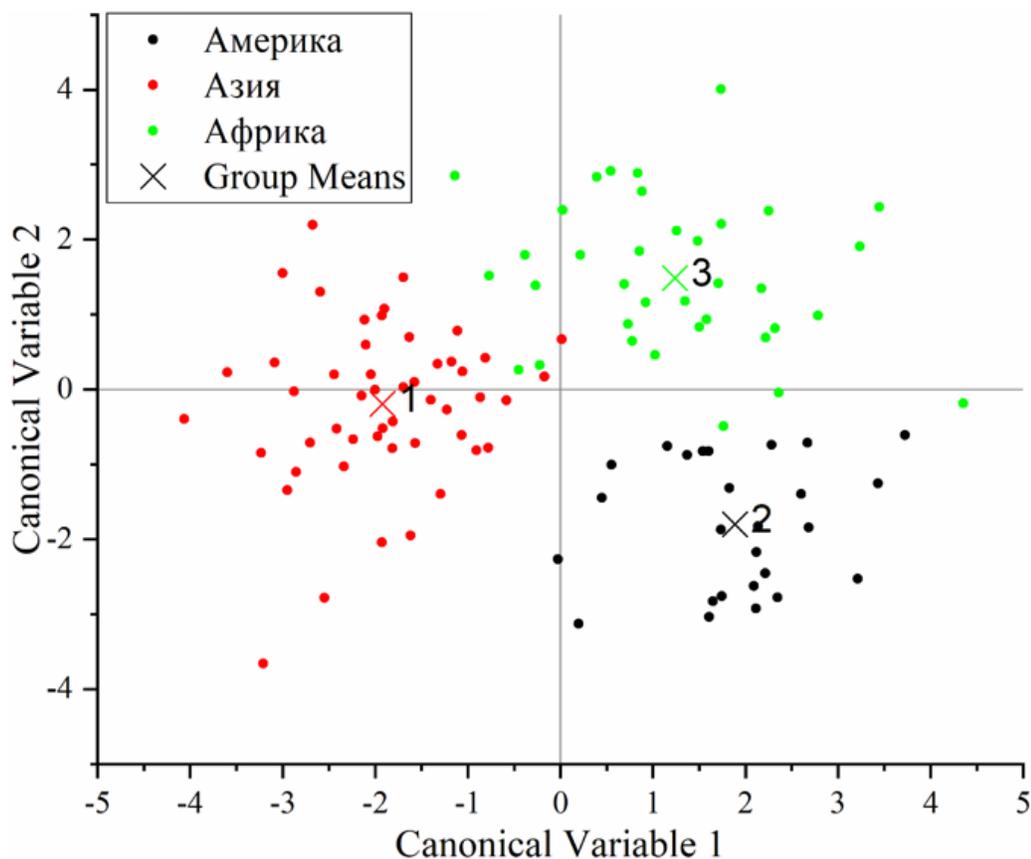


Рисунок 21 – Расположение точек, соответствующих образцам 1-3 групп кофе в координатах двух дискриминантных функций

Источник: составлено автором

В результате проведенных исследований были выявлены критерии идентификации гречневой крупы по времени сбора крупы (осень/весна), гидротермической обработке (ГТО) и срокам хранения по *UV-VIS-NIR*-спектрам и координатам спектроколориметрической системы *CIE L*a*b**; критерии идентификации – функциональные группы полос поглощения ИК-спектра в диапазоне $1700-700\text{ см}^{-1}$ для отличия зернового кофе от декофеинизированного кофе; разработана система классификационных функций градации образцов зернового кофе по географическому месту произрастания.

Глава 4 Разработка калибровочных моделей определения показателей качества пищевых продуктов сочетанием методов ИК-спектроскопии и многомерного анализа

4.1 Идентификация и подтверждение подлинности (аутентичности) сливочного масла при возможной фальсификации пальмовым маслом

В данном разделе диссертационного исследования представлено решение задачи ассортиментной и качественной идентификации и построения калибровочной модели определения содержания пальмового масла в составе сливочного и растительно-сливочного масла [18].

В качестве объектов исследования использовалось два набора образцов: калибровочный (обучающий) и тестовый (проверочный) для проверки адекватности хемометрической модели. Для формирования калибровочного набора образцов использовали пальмовое и сливочное масло (табл. 3).

ИК-спектры были объединены в матрицу X размерностью 75×1748 , где 75 – количество строк, которые соответствуют спектрам образцов, а 1748 – количество столбцов, соответствующих отсчетам на различных длинах волн (разрешающая способность спектрометра 2 см^{-1}). Для построения калибровочной модели использовалась база данных, состоящая из двух блоков: первый блок X – матрица ИК-спектров образцов из калибровочного и тестового наборов, второй блок Y – матрица содержания пальмового масла в составе смесей из калибровочного набора.

Анализ спектров образцов проведен в два этапа. На первом этапе проведено сравнение спектров и определены различия между образцами из калибровочного и тестового наборов. На втором этапе проведены построение калибровочной модели на основе хемометрических алгоритмов и проверка ее достоверности на тестовом наборе.

Спектры исследуемых образцов имели несущественные различия, а прямое сравнение спектров не позволяет осуществить ассортиментную и качественную идентификацию. В спектрах образцов масла преобладали типичные пики, которые могут быть отнесены к функциональным группам (жирные кислоты и эфиры). В связи с тем, что в сливочном масле присутствует большое количество функциональных групп (гидроксильные группы, кислоты, сложные эфиры, алифатические цепи жирных кислот и т.д.), это усложняет как процесс определения продукта по незначительным изменениям в спектрах, так и градацию образцов масла по составу.

Профили ИК-спектров калибровочного набора образцов смесей сливочного и пальмового масла, а также тестового набора образцов представлены на рисунке 22. Наиболее интенсивные пики полос спектра были обнаружены при 2928, 2851, 1740 и 1161 см^{-1} , которые связаны с наличием в образцах масла белковых и жировых связей [1,4,5,12,99,139]. Типичная спектральная область для жиров расположена в диапазоне от 1800 до 1400 см^{-1} , которая включает в себя валентные колебания С-Н, деформационные колебания С=О и С=C [8]. Пики абсорбции соответствовали различным конкретным диапазонам в видимой области ИК-спектра, а именно:

-2921 см^{-1} – валентные колебания С-Н в жирных кислотах;

-2851 см^{-1} – валентные колебания С-Н в жирных кислотах. Данный пик является типичным для пальмового масла, интенсивность поглощения растет пропорционально с увеличением процентного содержания пальмового масла в смеси [57];

-1740 см^{-1} – деформационные колебания С=О в кислотах и эфирах;

-1161 см^{-1} – валентные колебания С=О и С-С в кислотах;

-1650 см^{-1} – валентные колебания С=О в амидах.

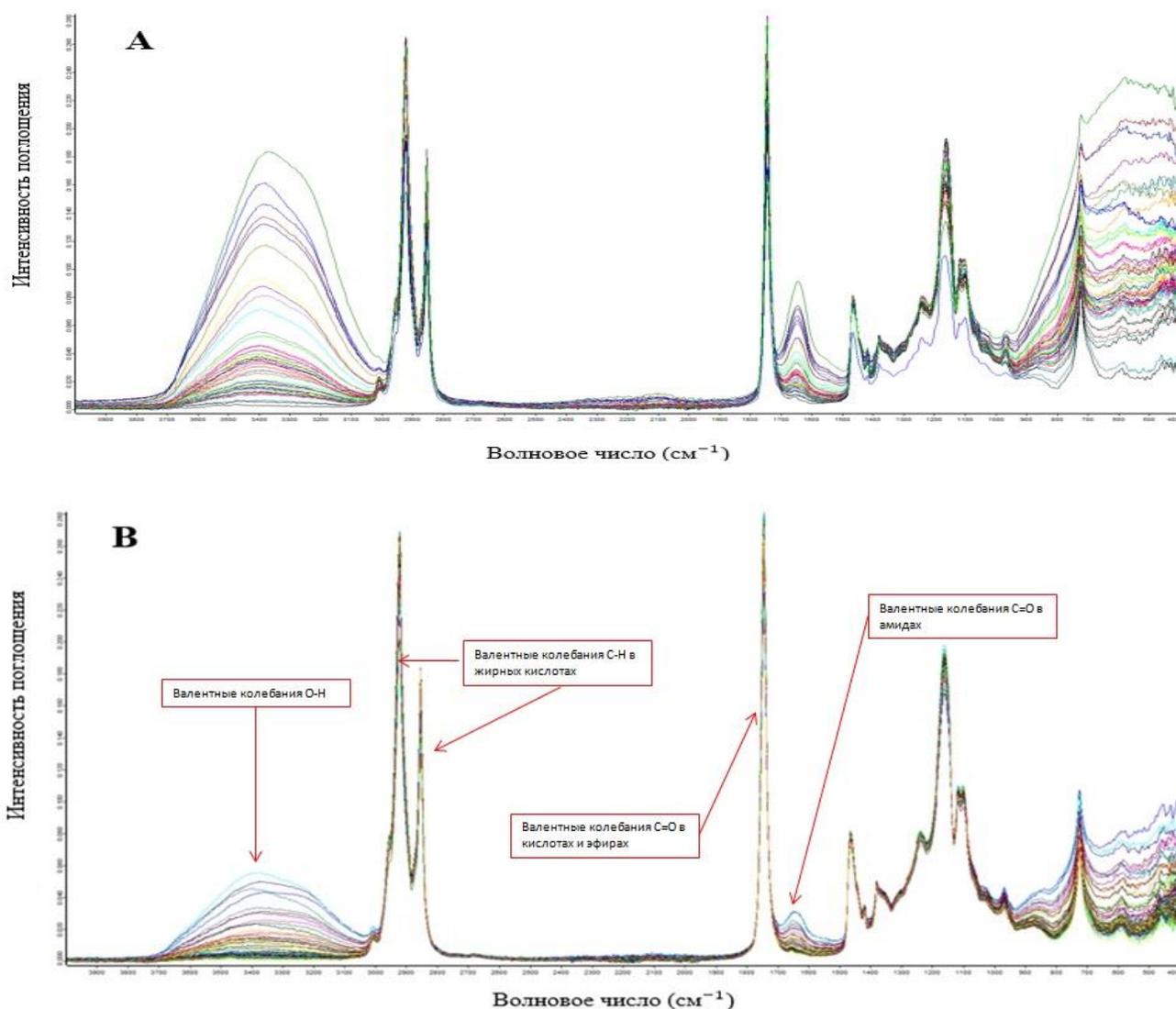


Рисунок 22 – Профили ИК-спектров калибровочного набора образцов смесей сливочного и пальмового масла (А), тестового набора образцов сливочного масла (В)
 Источник: составлено автором

Построение калибровочной модели проведено в несколько этапов:

- осуществлен выбор оптимального из двух методов многомерной градуировки: РГК и ПЛС;
- построены калибровочные модели ПЛС с различающимся спектральным интервалом и количеством факторов (латентных переменных) для увеличения точности итоговой модели;

-проведена проверка эффективности итоговой модели на образцах из тестового набора.

На этапе предварительного анализа был использован метод РГК для определения количества главных компонент и выбора спектрального интервала. При использовании метода РГК были определены области спектров образцов, которые вносят наибольший вклад, а также области с наименьшим вкладом, например, область от 4000 до 3110 см^{-1} (валентные колебания О-Н) была удалена при построении калибровочной модели. В результате проведенного анализа была получена оптимальная модель РГК и выявлена область спектра от 3030 до 2780 см^{-1} , которая определяет жировой состав образцов. Несмотря на это, полученная модель РГК имеет высокую среднеквадратичную ошибку валидации и прогноза, а при валидации модели на тестовом наборе образцов получены неудовлетворительные результаты. Поэтому модель РГК не может быть использована для достоверного определения фальсификации. Однако данные, полученные при применении метода РГК, позволили упростить построение калибровочных моделей методом ПЛС.

При использовании метода ПЛС проведено обнаружение выбросов в калибровочном наборе образцов по сочетанию F -статистики и критического уровня значимости статистики T^2 Хотеллинга. К выбросам, как правило, относятся те спектры, которые одновременно не соответствуют этим двум критериям. В наборе данных обнаружено 7 выбросов, которые не соответствуют только по одному из двух критериев, поэтому удаление их из модели нецелесообразно. Результаты проведенного моделирования представлены в табл. 9.

Таблица 9 – Проверка качества и точности калибровочных моделей идентификации сливочного масла

Метод	Спектральный интервал	Число факторов	Наименование выборки	<i>RMSEC/RMSEV</i>	R_c^2 / R_v^2
РГК	3030-2780 cm^{-1}	2	Обучающая	3,5149e-04	0,978
			Проверочная	4,3143e-04	0,967
ПЛС1	4000-400 cm^{-1}	5	Обучающая	0,018	0,986
			Проверочная	0,025	0,975
ПЛС2	3110-400 cm^{-1}	4	Обучающая	0,017	0,988
			Проверочная	0,022	0,981
ПЛС3	3110-2730, 1790-600 cm^{-1}	4	Обучающая	0,014	0,991
			Проверочная	0,017	0,989

Источник: составлено автором

Количество факторов определялось по графику объясненной дисперсии последовательно, до достижения высокой точности (*RMSE*) и качества модели (R^2). Использование в модели большего числа факторов может привести к переоценке, которая сопровождается моделированием шумов, а не содержательной информации. Наиболее репрезентативной моделью является модель ПЛС3, содержащая 4 фактора. Проверка качества калибровочной модели ПЛС3 проведена по коэффициенту детерминации ($R_v^2=0,989$), а точность модели – по среднеквадратичной ошибке валидации ($RMSEV=0,017$).

С помощью полученной калибровочной модели ПЛС3 осуществлен количественный анализ содержания пальмового масла в образцах из тестового (проверочного) набора ($n=13$) – рис. 23.

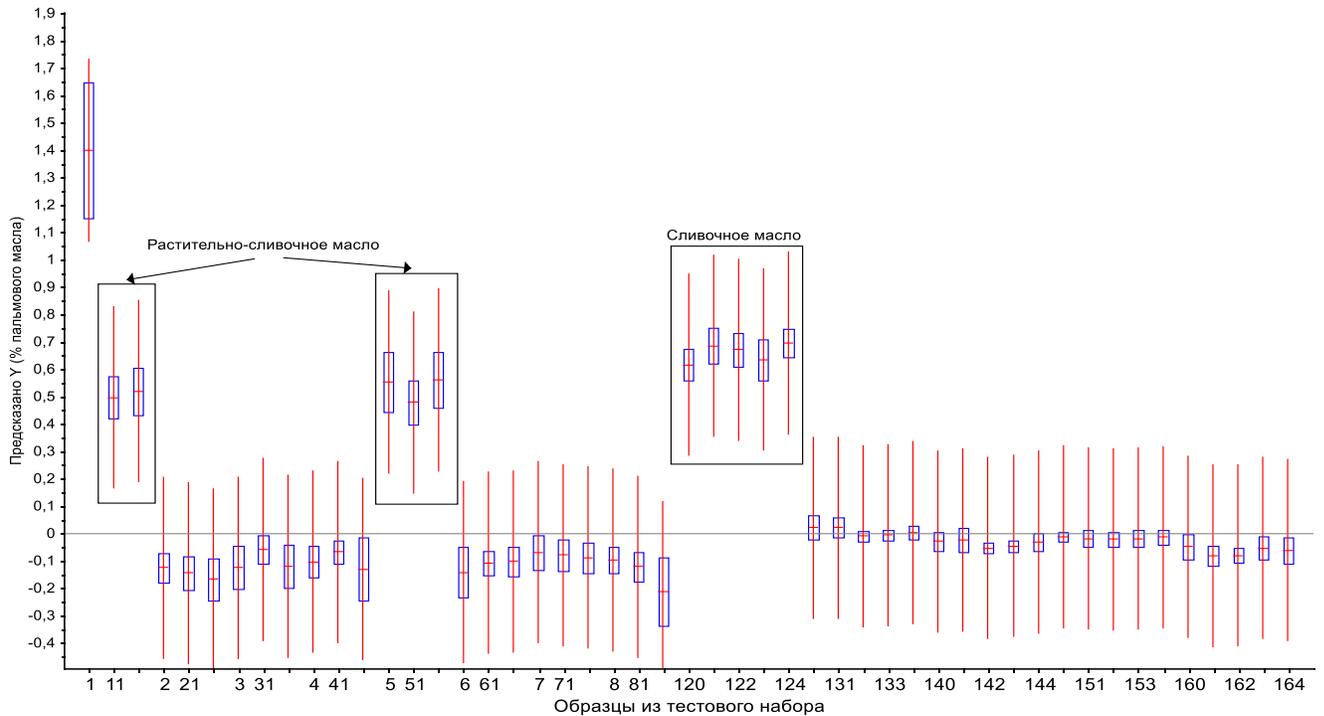


Рисунок 23 – Количественный анализ содержания пальмового масла в образцах сливочного масла из тестового набора по калибровочной модели ПЛС3

Источник: составлено автором

Проведена валидация калибровочной модели ПЛС3 по ИК-спектрам тестового набора образцов и выявлено: 10 образцов не содержат пальмового масла (сливочное масло); в 2 образцах обнаружено пальмовое масло (согласно маркировке образцов, они относятся к растительно-сливочному маслу); а в одном образце обнаружено пальмовое масло, что не указано на маркировке и соответственно является качественной фальсификацией. Количественное содержание обнаруженного пальмового масла в образцах из тестового набора представлено в табл. 10.

Таблица 10 – Количественное содержание пальмового масла в образцах из тестового набора

Наименование	Содержание пальмового масла, %	Отклонение, %
Образец растительно-сливочного масла №1	50	±7,5
Образец растительно-сливочного масла №2	52,3	±9,6
Фальсифицированный образец сливочного масла	65,4	±5,6

Источник: составлено автором

4.2 Прогнозирование кислотного числа жира и сроков хранения гречневой крупы

В качестве показателя качества (дескриптора) отобрано кислотное число жира (КЧЖ) гречневой крупы, как показатель свободных жирных кислот в продукте, напрямую зависящий от сроков хранения. Этот показатель наиболее интересен с потребительской точки зрения, так как позволяет сделать вывод о пригодности продукта для безопасного употребления. Результаты измерения КЧЖ образцов гречневой крупы из калибровочного набора представлены в таблице 11.

Таблица 11 – Результаты измерения КЧЖ образцов гречневой крупы из калибровочного набора

Наименование образца	КЧЖ, мг КОН/г (± δ)	Предел повторяемости r , %
«Винол» (длительного хранения)	13,13±1,57	0,53
«Скайфуд»	6,03±0,60	2,32
«Скайфуд» (длительного хранения)	10,6±1,27	4,70
«Бийский элеватор» (мелкая крупа)	14,08±1,68	1,00
«Бийский элеватор» (большая крупа)	12,41±1,48	5,10

Наименование образца	КЧЖ, мг КОН/г ($\pm\delta$)	Предел повторяемости r , %
«Зернышко к зернышку»	5,09 \pm 0,51	2,35
«Вкусвилл»	5,6 \pm 0,56	3,57

Источник: составлено автором

Спектральные данные экспортировали из встроенного программного обеспечения *OPUS 7.3.5.* в *Excel*, на основе которых создана матрица $X=J \times N$, где где 95 – количество строк, которые соответствуют спектрам образцов, а 1748 – количество столбцов, соответствующих отсчетам на различных длинах волн. Для построения калибровочной модели использовалась база данных, состоящая из двух блоков: первый блок X – матрица ИК-спектров образцов из калибровочного набора, второй блок Y – матрица значений КЧЖ и сроков хранения.

В процессе предварительного анализа построения калибровочной модели были поставлены и решены следующие задачи:

-осуществлен выбор оптимального метода многомерной калибровки: ПЛС. На этапе предварительного анализа также был использован метод РГК, но были получены неудовлетворительные результаты в связи с высокими значениями среднеквадратичной ошибки прогноза и валидации;

-построены калибровочные модели ПЛС с различающимся спектральным интервалом и количеством факторов (латентных переменных) для увеличения точности итоговой модели. Установлено, что для определения КЧЖ оптимален выбор спектрального диапазона 3700-2800 и 1800-700 см^{-1} , так как в указанном диапазоне проявляются колебания функциональных групп липидов (CH_2 , CH_3 , C=O) и OH -групп. Для построения калибровочной модели определения сроков хранения гречневой крупы выбран более узкий спектральный диапазон 3100-2800 и 1800-800 см^{-1} .

Проверка качества и точности калибровочной и прогностической моделей осуществлялась методом полной перекрестной валидации с использованием статистических показателей: R^2 , $RMSEC$, $RMSEV$. Результаты проведенного моделирования представлены в таблице 12.

Таким образом, использование контрольного анализа совместно с многомерными методами калибровки на основе данных ИК-спектроскопии позволяет определять КЧЖ и сроки хранения. Разработанные модели отличаются высокой точностью и качеством, поэтому могут использоваться для быстрого мониторинга качества гречневой крупы.

Таблица 12 - Проверка качества и точности калибровочных моделей идентификации гречневой крупы

Метод	Параметр	Спектральный интервал, см^{-1}	$RMSEC/$ $RMSEV$	$R_c^2/$ R_v^2
ПЛС	КЧЖ	4000-400	1,19/1,57	0,89/0,81
ПЛС	КЧЖ	3700-2800,1800-700	0,48/1,09	0,98/0,91
ПЛС	КЧЖ	3100-2800,1800-800	0,81/1,16	0,94/0,90
ПЛС	Срок хранения	4000-400	0,87/1,94	0,98/0,90
ПЛС	Срок хранения	3700-2800,1800-700	1,32/1,71	0,95/0,92
ПЛС	Срок хранения	3100-2800,1800-800	1,23/1,64	0,96/0,93

Источник: составлено автором

По результатам проведенных исследований были разработаны компьютерные модели по ИК-спектрам: определения содержания пальмового масла в составе сливочного и растительно-сливочного масла методом ПЛС; определения кислотного числа жира и сроков хранения гречневой крупы методом ПЛС.

Глава 5 Контроль качества пищевых продуктов методами анализа гиперспектральных изображений

5.1 Сортировка зерна гречихи с помощью гиперспектральных изображений

Для исследования использованы образцы зерна гречихи, описанные в табл. 3. Отобранное 81 зерно было подвергнуто рентгенографическому анализу на передвижной рентгенодиагностической установке ПРДУ-02 (ЗАО «ЭЛТЕХ-Мед», Россия), снимок (рис. 24) получен с помощью считывателя рентгеновского изображения *DIGORA* (*Soredex*, Финляндия) в лаборатории технологии длительного хранения продовольственных товаров и хлебопродуктов ФГБУ НИИПХ Росрезерва [3]. В состав установки входят: рентгенозащитная камера для проведения рентгенографических работ; источник рентгеновского излучения моноблочного типа РАП-70М; приемник рентгеновского изображения на основе экрана с фотостимулирующим люминофором (ФСЛ-экран); персональный компьютер с программным обеспечением.

Просмотр гиперспектрального изображения образцов зерна гречихи проводили с помощью программного комплекса Альбеда 4.0.23. С помощью функции отбора спектральных сигнатур были получены усредненные спектры 81 образцов зерна, которые были преобразованы в матрицу X размером 81×224 (81 – количество усредненных спектров образцов, 224 - отсчеты на длинах волн). Матрица X была использована для разработки модели классификации методом *PLS-DA* с помощью графического пользовательского интерфейса для программного комплекса *MATLAB 2018b* [188].

Для исследования зерен по признаку «невыполненность» из партии были отобраны 81 зерно гречихи. Зерна в количестве 81 зерно были приклеены на прямоугольную картонную рамку с липким верхним слоем (рис. 24). Скрытую дефектность зерен гречихи определяли по цифровым рентгеновским изображениям особенностей внутреннего строения и степени развития основных структур зерна (ГОСТ Р 59603-2021).

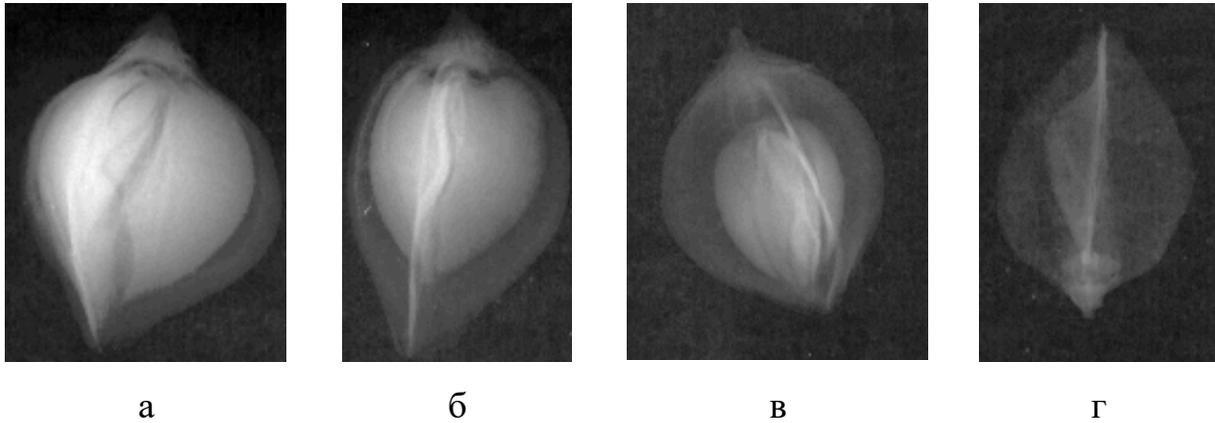


Рисунок 24 – Рентгенобразы образцов зерна гречихи различных групп по степени выполненности (а - 1 группа; б – 2 группа; в – 3 группа; г – 4 группа)

Источник: составлено автором

Анализ цифрового изображения зерна проведен на основе рентгенографических признаков; нерегулярность распределения оптической плотности по площади рентгеновской проекции, обусловленная развитием эндосперма, его отсутствием или неполным развитием.

Различные части зерна: семенная кожура, эндосперм, а также не- и поврежденные участки, ослабляют рентгеновское излучение в разной степени. Нормальные зерна – полностью сформированное без дефектов: эндосперм, область зародыша и оболочки равномерно светлые, без нерегулярных затемнений (рис. 24 а, б). Невыполненные зерна – в основном жизнеспособные. В отличие от выполненных зерен большая часть их проекции на рентгенограммах занята темным полем (рис. 24 в, г).

По результатам анализа на цифровом рентгеновском изображении зерна гречихи признак «невыполненности» выглядит как полное или частичное потемнение проекции в сравнении с изображением нормальных зерен, что указывает на более низкую плотность или толщину тканей зерна. В случае частичного потемнения видны очертания недоразвившегося зародыша, а при сплошном потемнении - зерно пустое, оно имеет незаполненную оболочку, а проекция зерна полностью темная.

Рентгеновские проекции зерна разделены на группы по степени выполненности, так и по соотношению участков «светлый – темный» рентгеновской проекции зерен следующим образом: 1 группа – выполненное зерно (ядро занимает 80-100% зерна); 2 группа – невыполненное зерно (ядро занимает 50-80% зерна); 3 группа – невыполненное зерно (ядро занимает 30-40% зерна/рудимент); 4 группа – пустые зерна (отсутствие ядра). По степени выполненности зерна из обучающей выборки распределились на группы следующим образом: в 1 группу было отобрано 35, во вторую – 32, в 3 группу – 9, в 4 группу – 3. Зерна гречихи из 1 и 2 групп может быть использовано для дальнейшей переработки и употребления в пищевых целях (доброкачественное), а зерно из 3, 4 групп является недоброкачественным.

Гиперспектральное изображение зерна гречихи (рис. 25) обрабатывали в программном комплексе Альбеда 4.0.23. С помощью встроенной функции отбора спектральных сигнатур на изображении были отобраны усредненные спектры каждого образца зерна ($n=81$) в диапазоне 935-1720 нм (рис. 26).

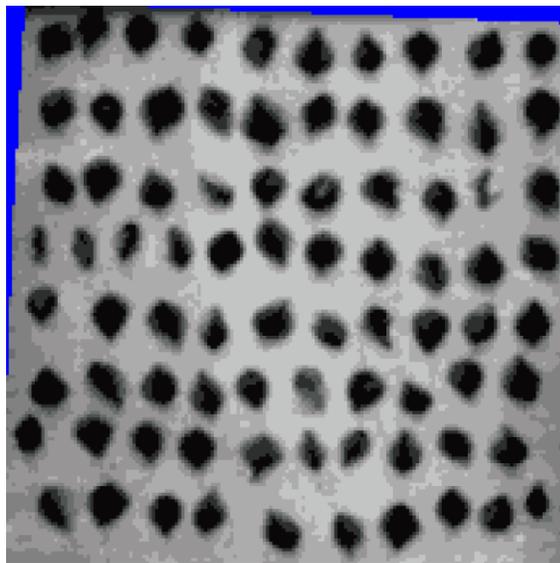


Рисунок 25 – Гиперспектральное изображение 81 образцов зерна гречихи в диапазоне 935-1720 нм

Источник: составлено автором

Наибольшие изменения в интенсивности отражения среди образцов различных групп отмечались при 1140 и 1440 нм. Образцам из групп 3,4 соответствуют большие значения интенсивности отражения при 973 нм (второй обертоном валентных колебаний ОН; вода), при 1140 нм (C=O; липиды) и при 1440 нм (первый обертоном валентных колебаний ОН; крахмал), что соответствует меньшему содержанию указанных компонентов в составе недоброкачественного зерна в отличие от доброкачественного [64,94]. Согласно полученным данным отобранные усредненные спектры образцов зерна гречихи чувствительны не только к составу шелухи, но и ядер гречихи.

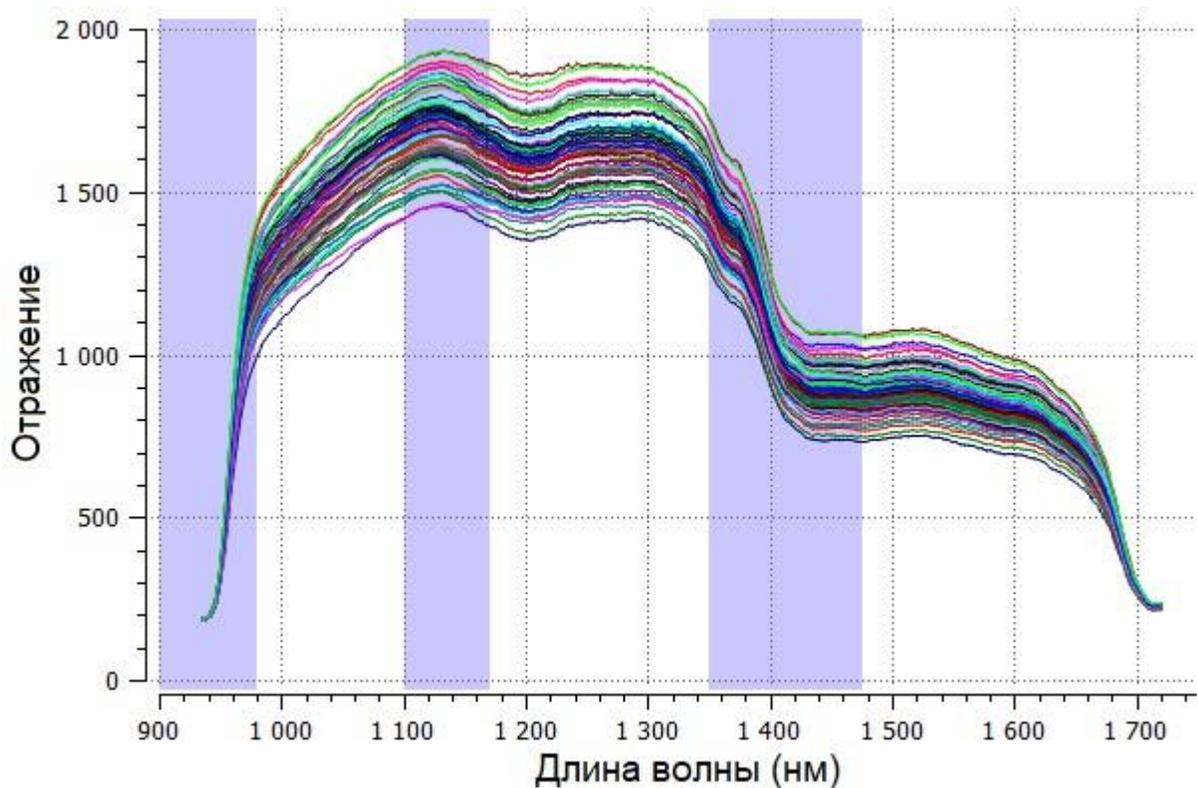


Рисунок 26 – Профиль спектров отражения в диапазоне 935-1720 нм образцов зерна гречихи

Источник: составлено автором

Матрица X усредненных спектров в диапазоне 935-1720 нм была использована для разработки модели классификации методом *PLS-DA*. Каждому образцу был присвоен класс – группа качества по степени выполненности зерна, согласно разделению

по данным рентгенанализа. После расчета матрицы «супер-счетов» была применена мягкая (*soft*) версия алгоритма *PLS-DA*, количество факторов – 12. Расположение координат объектов – образцов зерна гречихи разных групп - в пространстве двух плоскостей супер счетов представлено на рис. 27.

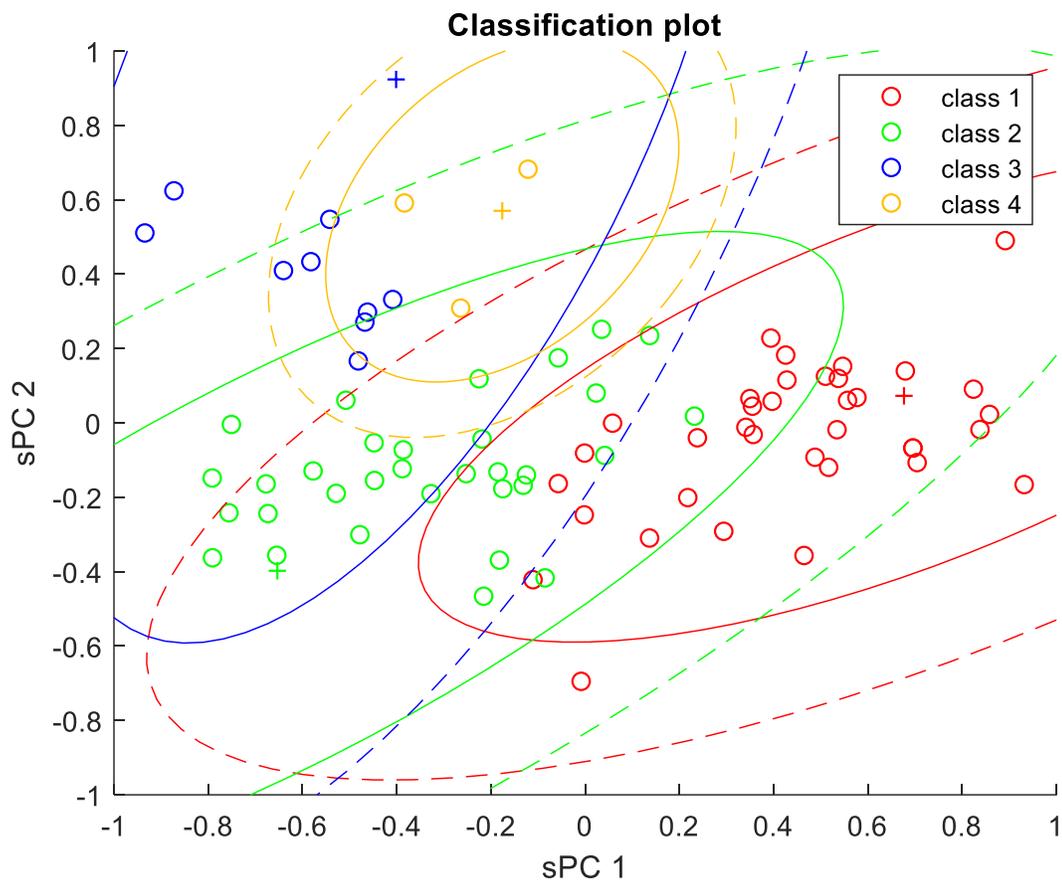


Рисунок 27 – Расположение точек, соответствующих спектрам образцов 1-4 групп зерна гречихи, в координатах двух плоскостей супер счетов модели *PLS-DA*

Источник: составлено автором

Оценка качества и точности полученной модели *PLS-DA* осуществлялась по статистическим параметрам: общая чувствительность (96%), общая специфичность (89%), общая эффективность (93%), а также при кросс-валидации методом K-fold с 10% удержанием набора валидации.

Внедрение гиперспектральных камер в ближнем инфракрасном диапазоне и сопутствующих хемометрических систем классификации зерна гречихи на категории качества имеет высокий потенциал для быстрого, неразрушающего и удаленного контроля качества. Использование таких систем обладает технико-экономической эффективностью и позволит уменьшить расходы по контролю качества на предварительных стадиях обработки зерна.

5.2 Классификация плодов авокадо Хасс по наличию дефектов с помощью гиперспектральных изображений

Стадии развития и созревания плодов авокадо важны для определения минимальных требований к качеству [90], которые позволяют продлить процесс созревания и достичь потребительской зрелости (*UNECE STANDARD FFV-42:2019*). Контроль качества плодов авокадо, включая проверку на зрелость, осуществляется как с помощью сенсорного анализа (*ISO 6658:2005 2017*), например, путем определения цвета кожицы (экзокарпа), так и инструментальными методами: измерение твердости плодов с помощью пенетрометра [145], измерение влажности, определение содержания липидов в мякоти (мезокарпе) [62]. Несмотря на свои очевидные преимущества такие методы являются разрушающими и трудоемкими.

HSI широко используется в промышленности, поскольку предоставляет большой объем количественной информации в видимой и инфракрасной области электромагнитного спектра, которую можно использовать для определения физико-химического состояния отдельного плода или урожая, мониторинга зрелости [107,154].

Образцы плодов авокадо после хранения были распределены между классами качества по наличию дефектов в соответствии с положениями стандарта *UNECE STANDARD FFV-42:2019*.

При анализе гиперспектральных изображений и разработке моделей классификации плодов авокадо на классы качества по наличию дефектов были использованы следующие многомерные методы: дисперсионный анализ (*ANOVA*) и МГК, *PLS-DA*, *SIMCA*, в программных приложениях *Albedo 4.0.23* (МФТИ, Россия) и *Prediktera Breeze ver. 2021.1.0* (*Prediktera AB*, Швеция). *PLS-DA* является общим хемометрическим методом для многомерной дискриминации и классификации данных [188], основанным на использовании регрессии *PLS2* для корреляции между матрицей предикторов X / матрицей откликов Y и фиктивными переменными. *SIMCA*, в свою очередь, также является методом классификации данных, основанным на анализе образцов из каждого класса с помощью метода главных компонент.

Классификационное моделирование включало сегментацию гиперспектральных изображений с использованием выделения *ROI* (весь плод, эллипсы), предварительную обработку спектральных данных: Стандартная нормальная вариация (*SNV*-коррекция) и центрирование. Центрирование - это вычитание среднего значения из каждой переменной для анализа дисперсии относительно среднего значения. *SNV*-коррекция - это преобразование, которое отделяет эффекты рассеяния света от спектральных данных путем центрирования и масштабирования отдельных спектров. При хемометрическом моделировании *SNV*-коррекция использовалась для уменьшения смещения фона в гиперспектральном изображении.

Сенсорный анализ экзокарпа плодов авокадо был проведен до хранения, после 10 дней хранения были исследованы экзо- и мезокарп образцов. Плоды авокадо сорта Хасс имели специфические внешние признаки: овальную форму, кожистый и буроватый экзокарп. Плоды были близкого размера, здоровые, чистые, в стадии потребительской зрелости, не перезрелые, без значительных механических повреждений, без избыточной внешней влажности, с аккуратно срезанной плодоножкой. В двух образцах плодов авокадо сорта Хасс (№ 4 и 5) были обнаружены незначительные поверхностные дефекты в виде морщин и пробкового покрытия на коже, соответственно.

Результаты сенсорного анализа состояния плодов авокадо до и после хранения представлены в таблице 13.

Таблица 13 – Состояние экзокарпия плодов авокадо до и после хранения

Номер образца	Описание состояния плодов авокадо по сенсорной оценке	
	До хранения	После хранения / категория качества
1	Дефекты отсутствуют	Локальное потемнение / <i>crit</i>
2	Дефекты отсутствуют	Локальное потемнение / <i>crit</i>
3	Дефекты отсутствуют	Повреждения, вызванные низкой температурой / <i>slight</i>
4	Дефект кожуры (морщины)	Дефект кожуры (морщины) / <i>slight</i>
5	Дефект кожуры (опробковение)	Дефект кожуры (опробковение) / <i>slight</i>
6	Дефекты отсутствуют	Дефекты отсутствуют / <i>good</i>
7	Дефекты отсутствуют	Дефект кожуры (морщины) / <i>slight</i>
8	Дефекты отсутствуют	Дефекты отсутствуют / <i>good</i>
9	Дефекты отсутствуют	Дефект кожуры (морщины) / <i>slight</i>
10	Дефекты отсутствуют	Дефекты отсутствуют / <i>good</i>

Источник: составлено автором

Сенсорный анализ состояния экзо- и мезокарпа, проведенный после хранения плодов авокадо, выявил следующее:

-образец плодов авокадо № 5 (рис. 28а) характеризовался опробковением экзокарпа, однако изменений в консистенции мезокарпа обнаружено не было. Данный дефект относится к допустимым, так как не приводит к критическим изменениям мякоти плода;

-образец плодов авокадо № 3 (рис. 28б) характеризовался локальным потемнением и размягчением мезокарпа. Обнаруженный дефект предположительно был вызван воздействием низкой температуры, но был допустимым, так как площадь повре-

жденной мякоти не превышала 6 см². Специфические признаки охлаждения проявляются не сразу, а через 12-48 часов, поэтому данный дефект может быть отнесен к скрытым;

-образцы № 4, 7 и 9 (рис. 28в) характеризовались дефектом кожицы в виде морщин. Исследование мезокарпа выявило незначительное размягчение мякоти, но изменения не носили прогрессирующего характера, что позволяет отнести их к допустимым дефектам;

-у двух образцов плодов авокадо (№ 1 и 2) в течение периода хранения развились критические дефекты (рис. 28г). На поверхности плодов авокадо появились округлые коричневато-черные вдавленные пятна, а мякоть под пораженным участком кожицы потемнела и размягчилась. Эти признаки могут быть связаны с грибковым заболеванием или перезреванием [109]. Микроорганизмы и грибы, присутствующие в воздухе и на поверхности плодов, могут проникать в ткани плодов через различные микроповреждения экзокарпа при сборе, упаковке и транспортировке плодов. Поскольку грибковое заболевание быстро передается здоровым плодам, важно обнаружить эти критические дефекты на ранней стадии, чтобы удалить поврежденные плоды из основной партии;

-образцы плодов авокадо №№ 6, 8 и 10 характеризовались отсутствием дефектов экзо- и мезокарпа. Развитие допустимых и критических дефектов плодов не регистрировалось. Сенсорный анализ плодов авокадо позволил отметить усиление стадии созревания, характеризующееся потемнением кожицы и уменьшением плотности мезокарпа.

По результатам сенсорного анализа плоды авокадо были разделены на три класса (табл. 13) в соответствии с *UNECE STANDARD FFV-42:2019*: критические дефекты - "*crit*"; незначительные поверхностные дефекты - "*slight*"; здоровые плоды, без дефектов - хорошее качество "*good*"). Согласно стандарту *UNECE STANDARD FFV-42:2019*, авокадо делятся на три класса: "Экстра" класс, I класс, II класс, классы которых, как правило, отличаются рядом незначительных поверхностных дефектов. Так,

для класса I площадь дефектов мезокарпа не должна превышать 4 см², а для класса II - 6 см². Таким образом, после сенсорного анализа плоды авокадо, имеющие недопустимые дефекты, были отнесены к отдельной категории *Crit* [129].



а)



б)



в)



г)

Рисунок 28 – Мезокарп образцов плодов авокадо с различными поверхностными дефектами кожицы: (а) № 5 с пробкой - без изменений мезокарпа; (б) № 3 с дефектами, возникшими в результате воздействия низкой температуры - локальное подрумянивание и изменение плотности мезокарпа; (в) № 4 с морщинами - незначительная потеря плотности мезокарпа; (г) № 2 с признаками заболевания (антракноз) - значительное потемнение и потеря плотности мякоти.

Источник: составлено автором

Во время хранения плодов авокадо в присутствии кислорода происходит ряд естественных окислительных процессов, сопровождающихся, помимо прочего, активацией ферментов. Изменение плотности мезокарпа авокадо связано с превращением протопектина в пектин, оно усиливается по мере созревания и сопровождается распадом растительной ткани мякоти и увеличением влажности [142]. В то же время значительное снижение плотности мякоти плода по отношению к экзокарпу приводит к образованию пустот под кожицей и более активному воздействию кислорода на мякоть авокадо, что ускоряет окисление. Согласно проведенным исследованиям [113] в переспелых плодах авокадо окисление приводит к критическим дефектам (значительное потемнение мякоти, ухудшение потребительских свойств). Несвоевременное обнаружение скрытых дефектов плодов авокадо, быстро развивающихся в период дозревания, может привести к снижению показателей качества всей партии плодов авокадо и увеличению потерь при хранении. Спецификация классов, установленных стандартом *UNECE STANDARD FFV-42:2019*, была использована для создания надежной и точной классификационной модели, которая позволит быстро определить плоды со скрытыми и непереносимыми дефектами и принять решение об оптимальном сроке хранения и периоде реализации плодов авокадо [129].

Моделирование классификации плодов авокадо на классы качества по наличию дефектов с использованием *HSI*-технологии включает несколько этапов: регистрация и анализ изображений методом *ANOVA* и МГК, сегментация изображений (выделение *ROI*), предварительная обработка (*SNV*-коррекция, центрирование), выбор метода многомерной классификации и спектрального диапазона, проверка модели.

Анализ гиперспектральных изображений методом ANOVA. Согласно оценке дисперсии данных гиперспектрального изображения, были выбраны спектральные длины волн с максимальной дисперсией при 550, 740 и 965 нм. Принципиальные различия в спектральных характеристиках плодов авокадо анализировались в пределах спектральных длин волн с максимальной дисперсией при 550, 740 и 965 нм:

При анализе гиперкуба данных оценивали вариацию данных, в результате чего были отобраны длины волн спектра с максимальной дисперсией при 550, 740 и 965 нм. Эти длины волн были использованы как зеленый (550 нм), красный (740 нм) и синий (965 нм) цвета, после чего слиты в одно *RGB*-изображение (рисунок 29). Основываясь на полученном *RGB*-изображении отбирали различные области интереса плодов авокадо для хемометрического анализа: целый плод, эллипсы.

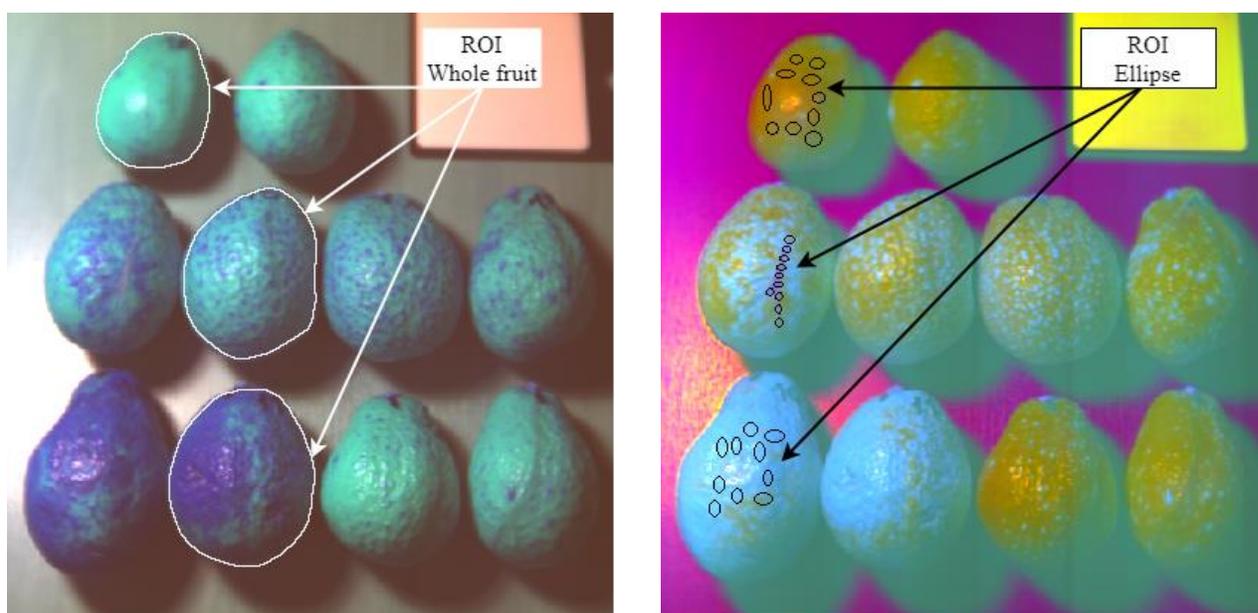
Основные различия в спектральных сигнатурах плодов авокадо обнаружены при 550, 740 и 965 нм:

-Полоса отражения при 550 нм соответствует отражению хлорофилла [119]. При сенсорном анализе плодов авокадо с признаками антракноза было зафиксировано значительное потемнение экзокарпа. Это связано с деградацией хлорофилла а до феофитина а [125], которая усиливается при бактериологическом поражении плодов и сопровождается, в частности, потемнением мезокарпа. Сравнение спектральных признаков поглощения показало смещение полосы поглощения хлорофилла а с 666 нм до 672 нм для образцов плодов авокадо с признаками антракноза.

-Красная область спектра, включающая диапазон длин волн от 690 до 750 нм [55], коррелирует с физиологическим состоянием растений. Согласно [58] область красного края чувствительна к изменениям в содержании хлорофилла, а максимальный наклон в области красного края свидетельствует о стадии спелости плодов. В исследовании [113] также отмечается, что отражение в области красного края постепенно уменьшается по мере созревания плодов авокадо и, в частности, перезревания. Красная область спектра тесно связана с наличием дефектов в плодах авокадо: максимальный наклон для образцов с плесневым повреждением экзокарпа и минимальный наклон для плодов без дефектов. Для плодов с дефектами из классов *Slight* и *Crit* качества БИК-спектр в диапазоне 700-970 нм имеет более низкие значения отражения, чем для неповрежденных плодов из класса *Good*.

-Спектральная область от 900 до 965 нм соответствует первому обертому частоты валентных колебаний ОН [94,145]. Созревание плодов авокадо сопровождается

уменьшением массовой доли воды при увеличении массовой доли липидов и сухого вещества, потемнением кожуры и появлением на плодах характерной для сорта окраски экзокарпа [47]. Возможно, различия в интенсивности отражения при 965 нм обусловлены процессами модификации хлорофилла и последующим изменением преломляющих свойств клеток, проявляющихся в ближнем инфракрасном диапазоне [55].



а)

б)

Рисунок 29 – RGB-изображение плодов авокадо с использованием в качестве цветов а) длин волн спектра; б) главных компонент. На рисунке а) показана сегментация гиперспектрального изображения путем выделения целого плода в качестве *ROI*, на рисунке б) *ROI* представляют собой эллипсы

Источник: составлено автором

Длины волн с максимальной дисперсией были отнесены к псевдоцветам и использованы для построения *RGB*-изображения (рис. 29а): 550 нм - зеленый, 740 нм - красный, 965 нм - синий. По *RGB*-изображению (рис. 29а) плоды авокадо разделены на классы качества по наличию дефектов: плоды слабого и критического классов ка-

чества имеют синий псевдоцвет (965 нм), плоды с антракнозом имеют более насыщенный синий цвет по всей поверхности; плоды хорошего класса качества имеют зеленый псевдоцвет (550 нм), что связано с отсутствием процессов деградации хлорофилла. *RGB*-изображение с использованием псевдоцветов позволяет оценить состояние плодов и подтвердить длины волн, которые в значительной степени отвечают за общую дисперсию гиперспектрального изображения, выбранную с помощью дисперсионного анализа *ANOVA*.

Анализ гиперспектрального изображения образцов плодов авокадо проводился методом главных компонент. В результате анализа были определены четыре главных компоненты: первая, вторая, третья и четвертая ГК в комбинации составляют 99,9% общей дисперсии (81,6%, 15,5%, 1,8% и 0,9%, соответственно) [129].

Тестовые коэффициенты, показывающие корреляцию коэффициентов отражения на длинах волн для 1-4 ГК, представлены на рис. 30. Эти данные были использованы для интерпретации ГК.

Первая ГК, объясняющая 81,6% общей дисперсии, связана со всеми коэффициентами отражения длин волн спектра, представлена со знаком (+) во всем спектральном диапазоне и, возможно, отвечает за общую дисперсию и корреляцию между пикселями (включая пиксели образцов авокадо) гиперспектрального изображения (рис. 30) [129].

Для второй ГК в соответствии со знаком (+/-) и тестовыми коэффициентами длин волн спектр делится на два диапазона: знак (+) относится к тестовым коэффициентам длин волн спектра, соответствующих *VIS*-диапазону, знак (-) - к *NIR*-диапазону (рис. 30). По аналогии с тестовыми коэффициентами по первой ГК, вторая ГК относится ко всем коэффициентам отражения длин волн спектра, но разделяет диапазон значений знаком (+/-) [129].

Для третьей ГК были обнаружены диапазоны полос, которые относят к каротиноидным пигментам (429, 434 и 458 нм) [86] и к 1-му обертону ОН (923 нм) со знаком (+), к хлорофиллу *a* (687 нм) и хлорофиллу *b* (592 нм и 603 нм) [51] со знаком (-).

Коэффициенты испытаний при 592 и 687 нм связаны с уменьшением содержания хлорофилла в плодах авокадо. Снижение концентрации хлорофилла в процессе хранения увеличивает содержание каротиноидных пигментов в экзокарпе плодов авокадо. Увеличение содержания лютеина (желтого пигмента, относящегося к классу каротиноидных пигментов) в кожице и мякоти плодов авокадо во время созревания подтверждается исследованиями [32]. Таким образом, интерпретация третьей ГК связана с изменением отражения в диапазоне полос хлорофилла и лютеина во время хранения внутри экзокарпа плодов авокадо без скрытых дефектов, при этом стадия созревания плодов не является критической [129].

Для четвертой ГК были обнаружены максимумы полос, которые относятся к антоцианам (545 нм) [127], каротиноидам (434 и 458 нм), красному максимуму (756 нм) со знаком (-) против 1-го обертона ОН (923 и 966 нм), хлорофиллу *a* (687 нм) и хлорофиллу *b* (592 нм) со знаком (+). Как показано в исследовании [32], содержание антоцианов в экзокарпе плодов авокадо увеличивается по мере созревания плодов авокадо сорта Хасс. В то же время полоса отражения при 545 нм со знаком (-), соответствующая отражению антоцианов, служит индикатором критического состояния созревания плодов авокадо и косвенным показателем бактериального повреждения плодов. Такие плоды авокадо необходимо быстро отсортировать и удалить из основной партии, чтобы снизить потери при хранении. В свою очередь, максимальное значение тестовых коэффициентов по 4ГК со знаком (-) при 750 нм соответствует красному краю спектральной области, что коррелирует с физиологическим состоянием растений. В связи с вышеизложенным, установлено, что 4ГК отвечает за процесс деградации плодов, учитывая экстремумы тестовых коэффициентов со знаком (-) при 545 и 750 нм [129].

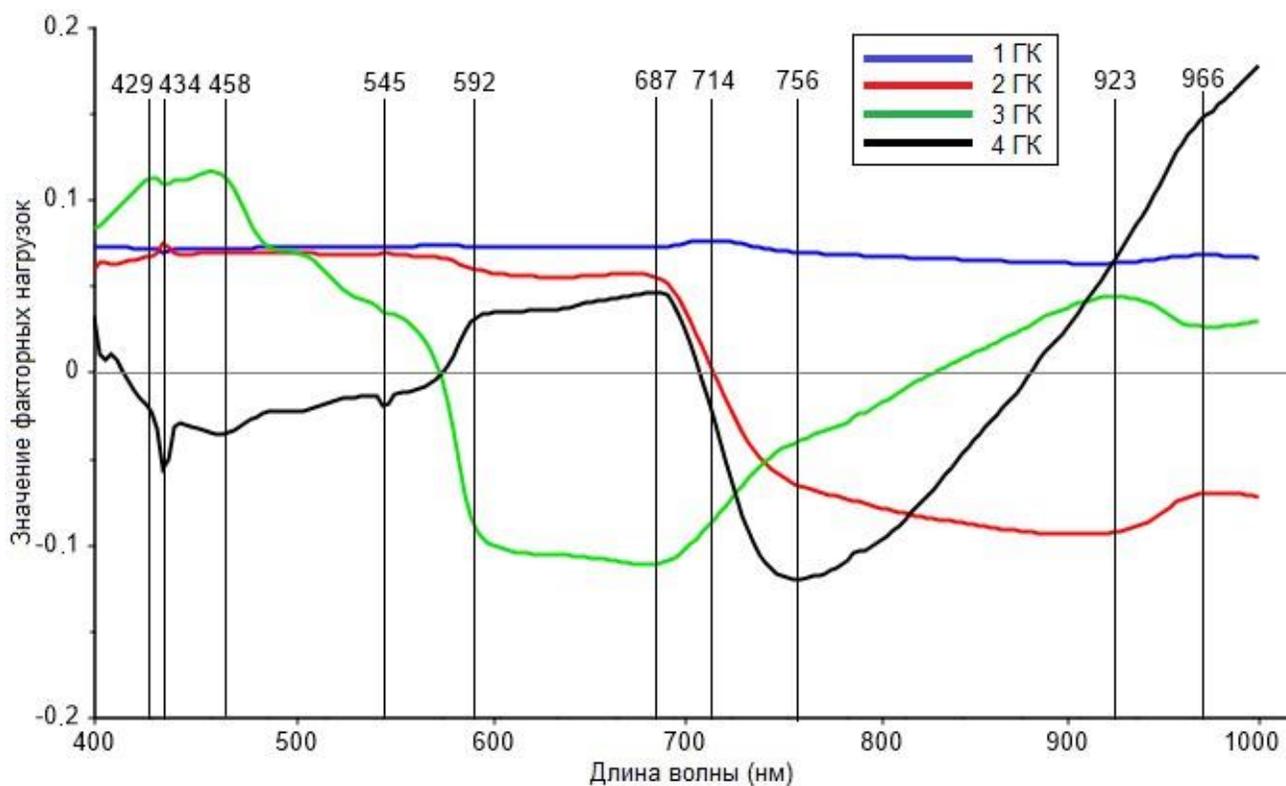


Рисунок 30 – Линейный график нагрузок МГК гиперспектрального изображения образцов плодов авокадо

Источник: составлено автором

В соответствии с интерпретацией ГК и их вкладом в общую дисперсию были выбраны три ГК, которые были отнесены к псевдоцветам (1-й ГК - красный, 3-й ГК - зеленый, 4-й ГК - синий), и было разработано следующее *RGB*-изображение (рис. 29б). Согласно рисунку (рис. 29б), 4ГК определяет дефектные плоды авокадо, которые представлены синим псевдоцветом.

RGB-изображения (рис. 29) с использованием псевдоцветов были использованы для сегментации *HSI* и разработки классификационной модели. На основе полученных *RGB*-изображений были выделены различные области интереса поверхности плода авокадо: целый плод, эллипсы в виде окружностей вокруг места дефекта.

При выборе *ROI* учитывались форма и интенсивность спектральных сигнатур плодов авокадо трех классов качества (рис. 29). Выбор *ROI* осуществлялся вручную

путем сегментации всего плода авокадо и выделения 10 спектральных сигнатур в виде эллипсов из каждого плода. При выделении эллипсов *ROI* гиперспектральное изображение увеличивалось, чтобы точно определить границы здоровой поверхности кожицы плодов авокадо и дефектной. После сегментации каждый *ROI* был отнесен к классификационным переменным - классам качества в зависимости от наличия дефектов. Каждый *ROI* как целого плода, так и эллипса представляет собой средний отклик спектральной сигнатуры плода авокадо одного из трех классов качества.

Классификация плодов авокадо на классы качества по наличию дефектов методами многомерного анализа проводилась с использованием различных методов многомерной классификации [166], областей интереса и спектральных диапазонов. Для повышения точности итоговых моделей была проведена калибровка с выделением областей интереса - эллипсов ($n=10$) на плодах авокадо разных классов качества. Преимущество выбора таких областей интереса заключается в том, что для каждого класса, определяемого по наличию дефектов, можно выбрать больше переменных, чтобы соответственно повысить воспроизводимость и адекватность предсказания.

В качестве предварительной обработки данных гиперспектрального изображения использовались *SNV* (*Standard Normal Variate*) -коррекция и центрирование. На этапе предварительного моделирования также использовалось сглаживание данных методом Савицкого-Голая, но полученные результаты по точности модели оказались неудовлетворительными (не показано в таблице).

Целые плоды ($n=5$) по всей площади были выбраны в качестве *ROI* для прогнозирования, что необходимо для корректного мониторинга качества после разработки модели. Результаты моделирования представлены в таблице 14.

Градация плодов авокадо на классы качества поддерживается сенсорным анализом, включая сенсорный анализ состояния мезокарпа после регистрации гиперспектральных изображений. Для разработки классификационной модели были использованы три класса качества в соответствии с сенсорным анализом: 1) *Crit* - плоды с вы-

раженным антракнозом кожицы (критические дефекты); *Slight* - плоды с повреждениями от низкой температуры, избыточной внешней влаги, пробковости кожицы и морщинистости кожицы (незначительные поверхностные дефекты); *Good* - плоды без дефектов.

При моделировании была зафиксирована низкая точность моделей с использованием в качестве *ROI* целых плодов, что связано с недостаточными для классификации наборами данных: при использовании *ROI* целых плодов количество переменных для калибровки составило 10, а при использовании *ROI* эллипсов - 100 (10 эллипсов от 10 плодов). В том числе низкая точность моделей с использованием *ROI* целых плодов обусловлена включением в *ROI* спектральных сигнатур теней и отражений, вызванных высокой освещенностью образцов галогеновой лампой. Зафиксировано также повышение точности моделей классификации в зависимости от выбора спектрального диапазона: исключение области 400-500 нм (слабый сигнал с низкой интенсивностью отражения или отсутствие сигнала), исключение области 990-1000 нм (значительные шумы). Таким образом, количество переменных и выбор спектрального диапазона напрямую влияет на точность моделей классификации.

Таблица 14 – Многомерные модели классификации авокадо Хасс по дефектам

Метод	<i>ROI</i>	Усредненный спектр			
		Спектральный диапазон, нм	Предварительная обработка (в порядке применения)	Точность калибровки, %	Точность при тест-валидации, %
<i>PLS-DA</i>	Целый плод	400-1000	<i>SNV</i> , центрирование	0	0
	Целый плод	500-990	<i>SNV</i> , центрирование	100,0	40,0
	Эллипс	400-1000	<i>SNV</i> , центрирование	100,0	60,0
	Эллипс	500-990	<i>SNV</i> , центрирование	100,0	80,0
<i>SIMCA</i>	Целый плод	400-1000	<i>SNV</i> , центрирование	36,4	20,0

Метод	ROI	Усредненный спектр			
		Спектральный диапазон, нм	Предварительная обработка (в порядке применения)	Точность калибровки, %	Точность при тест-валидации, %
	Целый плод	500-990	SNV, центрирование	36,4	20,0
	Эллипс	400-1000	SNV, центрирование	93,3	80,0
	Эллипс	500-990	SNV, центрирование	100,0	100,0

Источник: составлено автором

По результатам моделирования и тестовой проверки модель *SIMCA* с использованием эллипсов в качестве *ROI* в диапазоне 500-990 нм была выбрана как наиболее репрезентативная. При моделировании были выбраны критические расстояния (*Dcrit*) между классами. Диаграмма *Cootan's* модели *SIMCA* в рамках модельных координат *Good* (ось *X*) и *Crit* (ось *Y*) представлена на рис. 31. Более высокая точность и эффективность модели *SIMCA* по сравнению с моделью *PLS-DA* связана со структурой используемых данных [152]: в качестве целевого используется класс *Good*, а классы *Crit* и *Slight* можно рассматривать как один класс, включающий дефекты различной степени. При такой структуре данных подразумевается, что один класс находится в центре, а остальные классы занимают периферию [152]. Согласно [144,150] *SIMCA* является классификатором одного класса, поэтому в случае классификации плодов авокадо по наличию дефектов он является более эффективным алгоритмом по сравнению с *PLS-DA* [129].

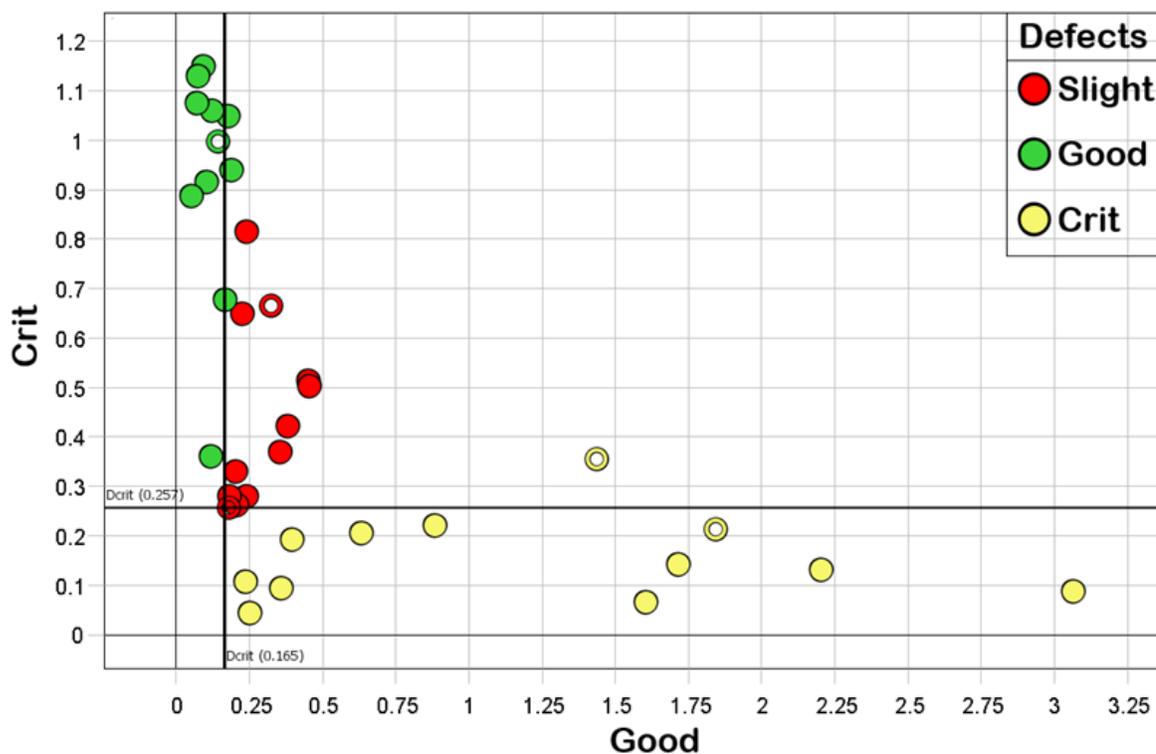


Рисунок 31 – График *Cootan's* модели *SIMCA* для классификации плодов авокадо по классам качества в зависимости от наличия дефектов

Источник: составлено автором

По завершении регистрации гиперспектральных изображений была построена многомерная модель *SIMCA* для классификации плодов авокадо по классам качества в зависимости от наличия дефектов. Качество и точность полученной модели подтверждены тест-валидацией с точностью 100% (табл. 14).

5.3 Сортировка плодов авокадо по гиперспектральным изображениям

Разработка технологии *HSI* позволяет проводить неразрушающую и бесконтактную цифровую сортировку плодов авокадо. Использование такого подхода дает возможность определять зрелость и проводить сортировку плодов в режиме реального времени и на месте. Поэтому актуальным и востребованным является разработка методических подходов построения калибровочных моделей определения $DM, \%$ и $MS, \%$

на основе *HSI* и методов многомерного анализа для определения зрелости и сортировки плодов авокадо [17].

Для исследования использованы калибровочный и тестовый набор образцов (табл. 3). Образцы плодов авокадо из тестового набора хранились в темном месте при комнатной температуре в течение 12 дней. Для образцов из тестового набора в течение сроков хранения отмечалось потемнение кожицы и снижение плотности мезокарпия, при этом указанные процессы ускорились на 7 день хранения. Начиная с 7 дня хранения по 12 день наблюдался прогрессивный характер возникновения критических дефектов: значительное потемнение кожицы, возникновение и распространение на ближайшие отдельные плоды грибковых заболеваний, а также значительное снижение плотности мезокарпия и возникновение внешней влажности у всех плодов из тестового набора (п. 5.2) [17].

Измерение влажности проводилось с использованием влагомера *Mettler Toledo NB43-S*. Для каждого образца авокадо из калибровочного набора отбиралась проба массой 10 г и высушивались при 100°C в до постоянной массы (45 мин). Разница между изначальной массой W_0 и массой после высушивания W_1 использовалась для расчета сухого вещества плодов авокадо из калибровочного набора, согласно уравнению (2).

Для плодов из калибровочного набора были получены гиперспектральные изображения ($n=4$) с каждой стороны (поворотом на 180°), что способствовало увеличению пространственных и спектральных данных при разработке калибровочных моделей. Регистрация гиперспектральных изображений образцов плодов авокадо из тестового набора проведена на 7,9 и 12 дни хранения [17].

При разработке калибровочных моделей определения $MS, \%$ и $DM, \%$ используют матрицы данных X и Y . Матрица X – это матрица предикторов размерностью $I \times J$, вторая матрица Y – матрица сигнатур в размерностью $I \times K$.

При построении калибровочных моделей определения $MS, \%$ и $DM, \%$ проводили отбор спектральных сигнатур (спектров в отдельных пикселях изображения) из гиперспектральных изображений плодов авокадо из калибровочного набора. По результатам отбора построена матрица данных X размерностью $I \times J$ (I – количество спектральных сигнатур HSI , J – число коэффициентов отражения на длинах волн) и матрица Y размерностью $I \times K$ (I – количество спектральных сигнатур HSI , K – число столбцов, равное числу значений показателей, характеризующих зрелость плодов ($MS, \%$ и $DM, \%$)).

Классификация образцов плодов авокадо из калибровочного набора посредством сегментации поверхности по HSI проведена методом главных компонент в программном комплексе Альбеда 4.0.23 (МФТИ, Россия).

Разработка калибровочных моделей осуществлялась в двух программных комплексах *Unscrambler X 10.0.4* (Camo Software, Норвегия) и *perClass Mira 3.1* (perClass BV, Нидерланды).

Качество и точность калибровочных моделей оценивалась по коэффициенту детерминации (R_C^2), среднеквадратичной ошибке ($RMSEC$) и по результатам полной перекрестной кросс-валидации моделей (R_P^2 , $RMSEV$, соответственно).

Влажность плодов авокадо. Зрелость авокадо тесно связана с содержанием влаги и сухого вещества плодов. Содержание сухого вещества плодов авокадо после срезания постепенно увеличивается [32]. Согласно стандарту *UNECE STANDARD FFV-42:2019* содержание сухого вещества плодов авокадо сорта Хасс в стадии потребительской зрелости должно быть не менее 21,0%. Перезревание, порча или хранение авокадо в ненадлежащих условиях, приводят к уменьшению содержания сухого вещества и увеличению влажности, что обусловлено окислением липидов мякоти и образованию свободной воды в составе плодов [17,149].

Таблица 15 – Содержание влажности и сухого вещества в образцах плодов авокадо из калибровочного набора

№ Образца	<i>MS, %</i>	<i>DM, %</i>
1	79,0	21,0
2	75,2	24,8
3	76,1	23,9
4	74,4	25,6
5	83,0	17,0
6	82,6	17,4
7	78,3	21,7
8	80,8	19,2
9	79,1	20,9

Источник: составлено автором

Из данных таблицы 15 следует, что диапазон значений влажности для образцов авокадо из калибровочного набора составляет 74,4-83,0 %, а сухого вещества путем расчета по значениям влажности – 17,0-25,6%, соответственно. Среди отобранных образцов, образцы под номером 5,6,8 с содержанием сухого вещества 17,0-19,2% не соответствуют требованиям стандарта, образцы плодов авокадо под номерами 2-4 имеют значение содержания сухого вещества выше минимальных требований стандарта (23,9-25,6% против 21,0%), что может свидетельствовать об их потребительской зрелости.

Показатели содержания влажности и сухого вещества использовались в качестве откликов у матрицы **Y** для разработки калибровочных моделей.

Содержание *MS, %* и *DM, %* сильно коррелирует с физиологическим состоянием плодов, включая, например, твердость плода. Согласно [83,115] физиологическое состояние плодов неоднородно как внутри единичной партии, так и по различным сортам авокадо. Согласно международному стандарту незначительный объем выборки

для проверки качества плодов авокадо не позволяет оценить множество отдельных плодов из партии. Цифровая сортировка по *HSI* позволит осуществлять градацию множества плодов авокадо по зрелости целой партии быстро и на месте [17].

На рисунке 32 представлены усредненные спектральные сигнатуры двух плодов авокадо из калибровочного набора, имеющие минимальное (образец №4) и максимальное (образец №5) содержание влаги. Как видно, в диапазоне 400-690 нм в *VIS*-области их спектры визуально слабо различимы, но значительно различаются в диапазоне *red-edge* (690-780 нм) *VIS*-области и в *NIR*-области (780-1000 нм). В связи с малым различием в *VIS*-области эти два образца имеют, соответственно, малое цветовое различие по окраске (ниже порога различения при визуальном осмотре).

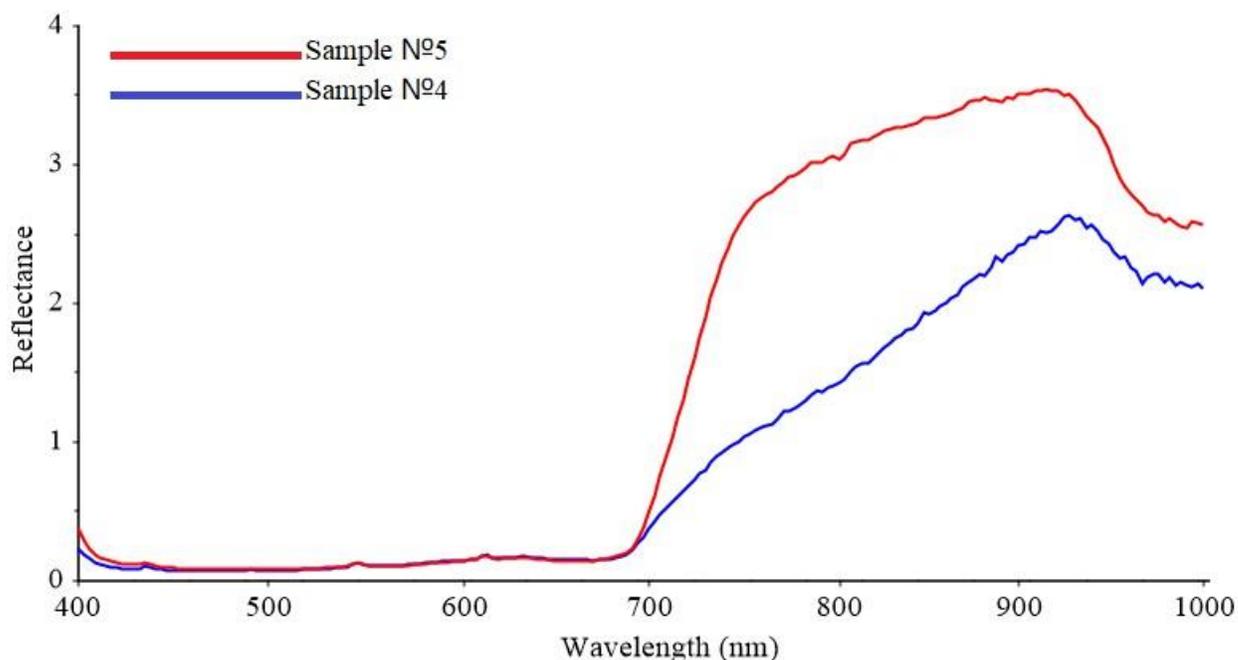


Рисунок 32 – Усредненные спектральные сигнатуры плодов авокадо из калибровочного набора: образец №5 (красная линия) и образец №4 (синяя линия)

Источник: составлено автором

Согласно опубликованным работам [45,92,153] в *NIR*-области присутствуют полосы колебаний С-Н при 740 нм [153] и широкая полоса третьего обертона колебаний

C-H, CH₂, и CH₃ при 840 нм [45,92], интенсивность которых коррелирует с содержанием сухого вещества. Следовательно, более низкое отражение в *NIR*-области при 740 и 840 нм характерное для образца №4, соответствует меньшему содержанию влаги и большему содержанию сухого вещества [17].

В качестве первого этапа разработки определения зрелости плодов авокадо необходимо изучить возможность сегментации поверхности по данным *HSI* в диапазоне 400-1000 нм. Такой подход, в свою очередь, позволит установить можно ли по *HSI* поверхности плодов авокадо: по отдельным сегментам поверхности или по всей площади поверхности производить их сортировку [17].

При анализе *HSI* с помощью МГК проведена декомпозиция трехмерной матрицы данных (512×512 *pixel*, 204 канала) с использованием алгоритма *NIPALS* на матрицы *T*-счетов и *P*-нагрузок. В матрице *T*-счетов приводятся пространственные координаты, т.е. пиксели изображения как независимые друг от друга данные. А в качестве *P*-нагрузок коэффициенты отражения спектра *HSI*.

По результатам обработки матрицы данных выделено первые четыре главные компоненты: первая, вторая, третья и четвертая ГК, которые объясняют 99,5% от общей дисперсии (70,3%, 27,0%, 1,4%, 0,8% соответственно). Значение факторных нагрузок коэффициентов отражения на длинах волн спектра по четырем ГК представлена на рисунке 33. Проведена интерпретация с первой по четвертую ГК по знаку (+/-) и по значению факторных нагрузок коэффициентов отражения на длинах волн спектра.

Первая ГК связана со знаком (+) с коэффициентами отражения спектра без явных максимумов. Вторая ГК согласно знаку (+/-) факторных нагрузок коэффициентов отражения на длинах волн спектра разделяет спектр на два диапазона: со знаком (+) – для факторных нагрузок в *VIS*-диапазоне от 400 до 710 нм, со знаком (-) – в *NIR*-диапазоне от 710 до 1000 нм. Следовательно, 1 и 2 ГК, во-первых, содержат большую часть информации (97,3% от общей дисперсии), содержащуюся в гиперкубе данных, но при этом, во-вторых, коэффициенты отражения на длинах волн спектра по этим

двум ГК сильно коррелированы и поэтому не вносят вклад, позволяющий проводить сегментацию поверхности, что показывает анализ изображений по этим ГК [17].

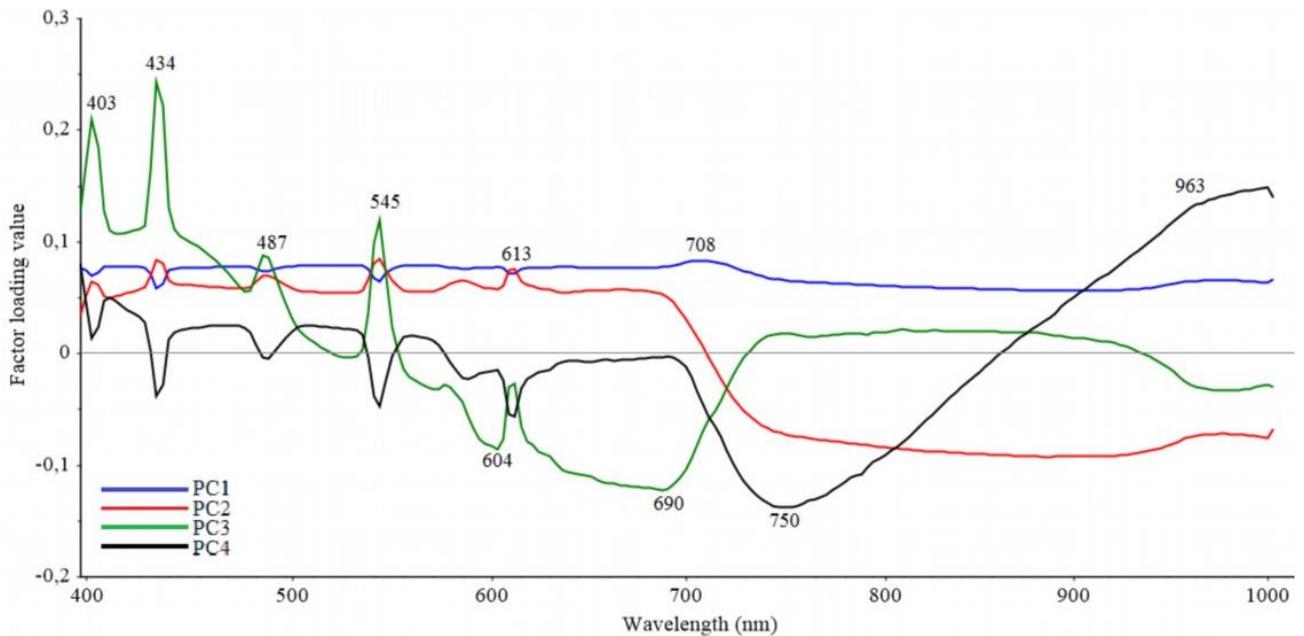


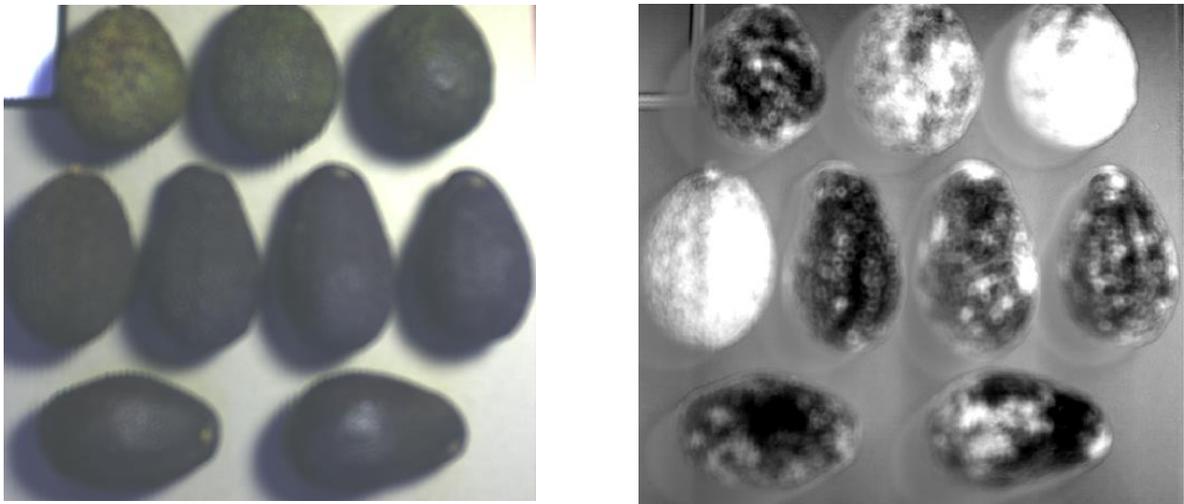
Рисунок 33 – Факторные нагрузки коэффициентов отражения на длинах волн спектра *HSI* в диапазоне 400-1000 нм по четырем ГК (PC1-PC4).

Источник: составлено автором

При анализе факторных нагрузок коэффициентов отражения на длинах волн спектра по 3 ГК выявлены следующие максимумы полос: 403, 434, 487, 545, 604, 690, а также широкая полоса от 740 до 940 нм. Ряд максимумов факторных нагрузок коэффициентов отражения на длинах волн спектра, соответствуют полосам поглощения пигментов в составе кожуры плода, а полоса в диапазоне 740 до 940 соответствует *red-edge*, что позволяет судить, что эта ГК соответствует степени потребительской зрелости плодов авокадо, но не описывает влажность плодов. Так, например, по 3 ГК присутствуют полосы отражения каротиноидов при 434 нм [86], антоцианов при 545 нм [32], полосы отражения хлорофилла *b* при 604 нм и хлорофилла *a* при 690 нм [55]. Знак (+/-) максимумов факторных нагрузок по 3 ГК, вероятно, связан с процессами изменения содержания пигментов в кожуре авокадо при созревании. Как показано в

работах [32,52] при созревании в кожуре авокадо наблюдается уменьшение общего содержания хлорофилла и каротиноидов при увеличении содержания антоцианов. При этом в первые дни созревания уменьшение содержания каротиноидов наблюдается незначительно [32]. Следовательно, максимумы факторных нагрузок при 434, 545 нм со знаком (+) связаны с изменением содержания антоцианов и незначительно с содержанием каротиноидов, а максимумы при 604 и 690 нм со знаком (-) – с содержанием хлорофилла при созревании плодов авокадо. Поэтому, 3 ГК информирует о соотношении содержания пигментов в составе кожуры при созревании плодов авокадо.

По 4 ГК со знаком (+) выявлены имеют факторные нагрузки, соответствующие полосе отражения 1-го обертона *ОН* при 963 нм [94,145], а со знаком (-) - спектральной области *red-edge* с максимумом при 750 нм.



а)

б)

Рисунок 34 – Гиперспектральное изображение плодов авокадо из калибровочного набора: исходное (а), в оттенках серого образцов с использованием 4 ГК как псевдо-цвета (б). Образцы авокадо расположены по порядковым номерам слева-направо и сверху-вниз

Источник: составлено автором

После вычисления ГК пользовательский интерфейс Альбеда 4.0.23. позволяет просматривать изображение при использовании ГК в качестве псевдоцветов в различных режимах отображения (оттенки серого, *RGB*). Для визуализации различий плодов авокадо из калибровочного набора по влажности 4 ГК была использована в качестве псевдоцвета для перестроения гиперспектрального изображения (рис. 34б) и его сравнения с исходным изображением (рис. 34а).

На рисунке 34а представлено изображение плодов авокадо снятых с помощью *RGB*-камеры по сравнению с изображением, построенного по значениям факторных нагрузок по 4 ГК (рис. 34б), значение которых представлено в оттенках от серого до черного цвета. При визуальном анализе изображения плодов авокадо снятых с помощью *RGB*-камеры не видно явных различий между образцами.

Из рисунка 34б видно, что образцы № 2,3,4 из калибровочного набора окрашены в светло-серые цвета, в отличие от других образцов, цвет площади поверхности которых ближе к черному цвету. Согласно данным таблицы 15, образцы № 2,3,4 более низкое значение влажности (от 74.4% до 76.1%) по сравнению с другими образцами (от 78.3% до 83.0%). Следовательно, по значениям факторных нагрузок по 4 ГК возможно проводить сортировку плодов авокадо по влажности, а также выявлено, что образцы неоднородны по спектральным сигнатурам не только между образцами из одной партии, но и по поверхности одного плода.

Разработка калибровочных моделей определения MS, % и DM, % по HSI. По результатам анализа *HSI* выбрано два направления разработки калибровочных моделей:

-посредством отбора спектральных сигнатур *HSI* случайным образом с поверхности плодов;

-использование всей площади поверхности *HSI* (исходные *HSI*) плодов авокадо в качестве *ROI*.

Разработка калибровочных по спектральным сигнатурам проведена методом ПЛС в программном комплексе *Unscrambler X 10.0.4*, по исходным *HSI* – в *perClass Mira 3.1 (perClass BV, Netherlands)*.

При разработке калибровочных моделей определения $MS, \%$ и $DM, \%$ плодов авокадо по спектральным сигнатурам проводили моделирование с изменением следующих параметров (табл. 16):

-использование спектральных сигнатур с гиперспектрального изображения с одной или с двух сторон;

-уменьшение количества использованных спектральных сигнатур при последовательном удалении выбросов согласно статистике *Hotteling's T²* и *F-residuals*;

-определение количества латентных переменных (факторов) по кумулятивной доле объяснимой дисперсии.

Таблица 16 – Проверка точности и качества калибровочных моделей с использованием спектральных сигнатур для определения сухого вещества и влажности плодов авокадо Хасс

Показатель, %	Количество использованных <i>HSI</i>	Количество спектральных сигнатур	Количество факторов	<i>RMSEC</i>	<i>RMSEV</i>	R_C^2	R_v^2
<i>MS</i>	1	185	7	1,46	1,64	0,79	0,73
<i>MS</i>	1	173	8	1,31	1,58	0,84	0,76
<i>MS</i>	1	127	8	0,87	1,11	0,93	0,89
<i>MS</i>	1	123	8	0,84	1,10	0,93	0,89
<i>MS</i>	2	410	10	1,44	1,57	0,77	0,74
<i>MS</i>	2	369	9	1,41	1,61	0,78	0,72
<i>MS</i>	2	296	11	0,99	1,10	0,89	0,87
<i>MS</i>	2	264	10	0,94	1,00	0,90	0,89
<i>DM</i>	1	185	8	1,33	1,54	0,82	0,76
<i>DM</i>	1	125	8	0,8	0,97	0,93	0,9
<i>DM</i>	1	103	8	0,59	0,77	0,97	0,94
<i>DM</i>	2	410	10	1,44	1,57	0,77	0,74
<i>DM</i>	2	309	10	0,92	0,99	0,91	0,90

Показатель, %	Количество использованных <i>HSI</i>	Количество спектральных сигнатур	Количество факторов	<i>RMSEC</i>	<i>RMSEV</i>	R_C^2	R_V^2
<i>DM</i>	2	257	9	0,78	0,86	0,94	0,92

Источник: составлено автором

По итогам моделирования (табл. 16) наиболее репрезентативными были отобраны модели, в которых используются спектральные сигнатуры гиперспектральных изображений с двух сторон плодов авокадо (рис. 35).

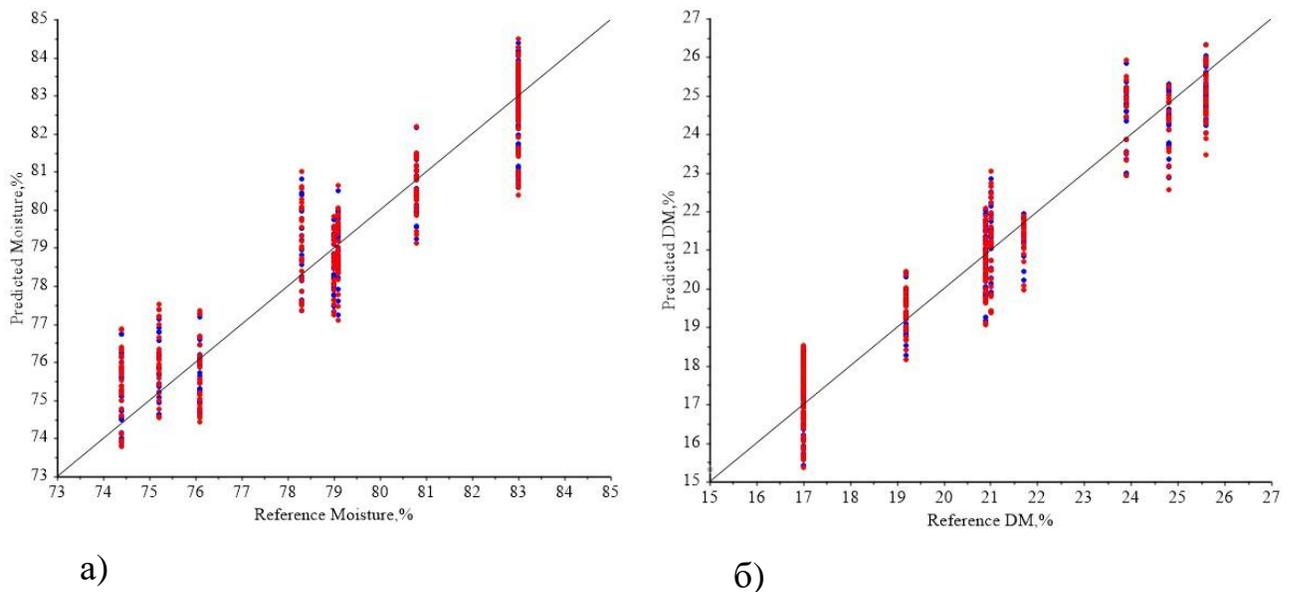


Рисунок 35 – ПЛС модели определения влажности (а) и сухого вещества (б) для плодов авокадо Хасс с использованием спектральных сигнатур

Источник: составлено автором

Низкие значения показателей R_C^2 и R_V^2 для таких моделей, в отличие от моделей со спектральными сигнатурами гиперспектрального изображения, с одной стороны, компенсируются меньшими значениями *RMSEC* и *RMSEV*, что повышает качество моделей. Последовательное обнаружение и удаление выбросов, то есть спектральных сигнатур, превышающих критический предел согласно статистике *Hotteling's T²* и *F*-

residuals, повышало качество и точность моделей. К выбросам были отнесены спектральные сигнатуры, отобранные с поверхности плодов авокадо в местах возникновения бликов от галогеновой лампы. Показано, что использование данных спектральных сигнатур гиперспектральных изображений с двух сторон позволяет увеличить качество моделей при незначительном снижении точности. Однако, отбор спектральных сигнатур вручную не позволяет в достаточной мере избавиться от фоновых шумов и выбросов в спектральных данных, содержащихся в областях с бликами от галогеновой лампы, что приводит к снижению точности и необходимости последовательного удаления выбросов из калибровки.

Разработку калибровочных моделей по исходным *HSI* плодов авокадо проводили с реализацией алгоритма: выделение плодов авокадо по контуру, удаление заднего фона и бликов от галогеновой лампы. После этого осуществлена сегментация объектов, то есть выделение каждого плода авокадо на передний план как отдельную область интереса (*ROI*).

Точность калибровочных моделей для определения *DM, %* и *MS, %* до удаления областей с бликами от галогеновой лампы составляла $R_C^2=0.93$ и $R_C^2=0.86$, соответственно. После сегментации *HSI* были разработаны калибровочные модели с высокой точностью $R_C^2=0.99$ для определения содержания *MS, %* (Модель 2) (рис. 36а) и *DM, %* (рис. 36б). Высокая точность достигается за счет увеличения количества репрезентативных спектральных данных.

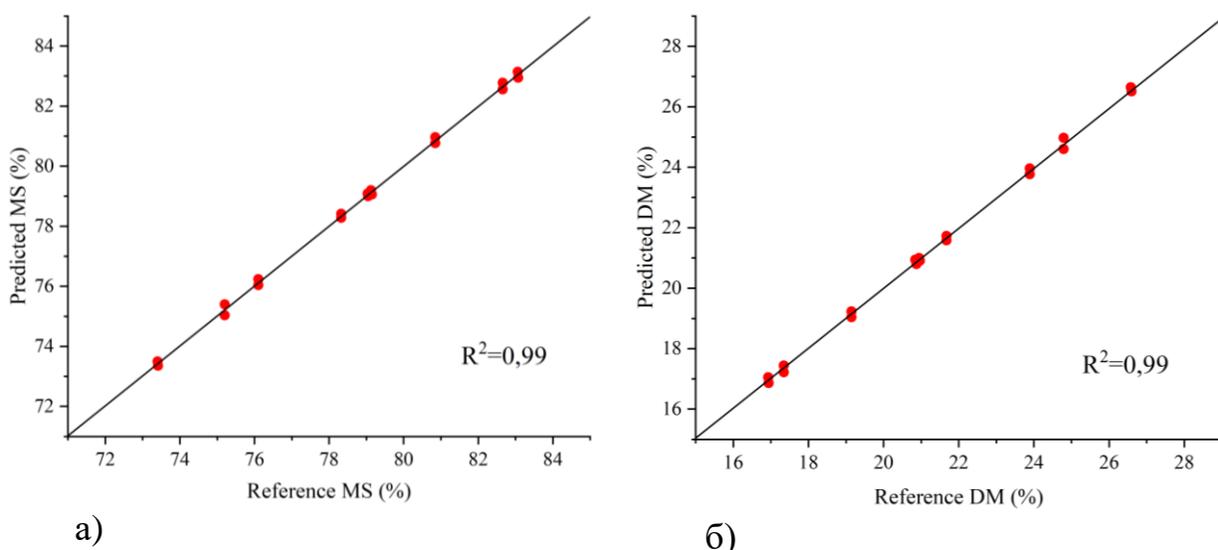


Рисунок 36 – Калибровочные модели определения влажности (а) и сухого вещества (б) для плодов авокадо Хасс с использованием исходных *HSI*.

Источник: составлено автором

В отличие от калибровочных моделей с использованием спектральных сигнатур, выбранных случайным образом с поверхности плодов, более высокая точность моделей с использованием всей площади поверхности *HSI* достигается за счет увеличения спектральных данных - числа пикселей с площади изображения каждого плода авокадо из калибровочного набора.

Тест-валидация калибровочных моделей проведена на тестовом наборе образцов плодов авокадо ($n=10$) со сроком хранения, равным 7,9 и 12 дней. Выбор тестовых образцов со сроком хранения превышающим установленные производителем сроки годности обусловлен необходимостью проверки двух разработанных калибровочных моделей. Для определения значения *MS, %* из тестового набора использованы Модель 1 (табл. 16) и Модель 2 (рис. 36а). Значения показателей *MS, %* плодов авокадо из тестового набора, определенные по одной из двух калибровочных моделей, усреднены по дням хранения (рис. 37).

Как видно из рисунка 37, значения показателей *MS, %* определенные по спектральным сигнатурам *HSI* выбранных случайным образом с поверхности плодов (Модель 1) незначимо различаются от срока хранения и соответствуют потребительской

зрелости плодов, что не согласуется с органолептической оценкой плодов из тестового набора. Модель 1 не позволяет принять достоверное решение по возможности употребления плодов вне установленных сроков хранения.

Значения $MS, \%$ определенные по исходным HSI с последующей обработкой (Модель 2) значительно различаются от срока хранения плодов, не соответствуют нормам стандарта *UNECE STANDARD FFV-42:2019* и в то же время хорошо согласуются с органолептической оценкой плодов из тестового набора.

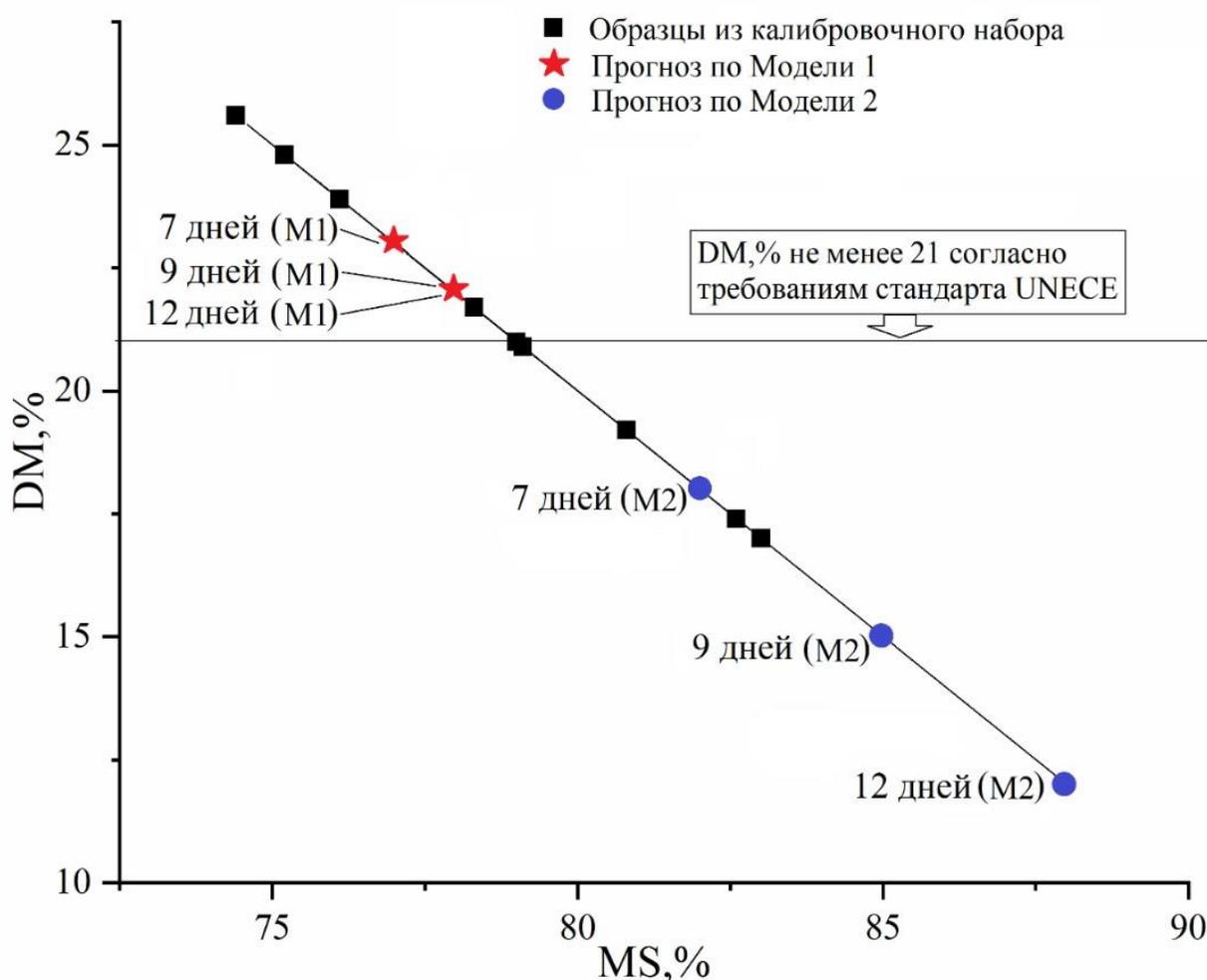


Рисунок 37 – Сравнение результатов тест-валидации Модели 1 и Модели 2 для определения $MS, \%$ образцов плодов авокадо из тестового набора от сроков хранения.

Источник: составлено автором

У всех плодов из тестового набора согласно органолептической оценке наблюдалось возникновение критических дефектов, снижение твердости плодов и возникновение внешней влажности, что подтверждается значениями $MS, \%$ определенными по Модели 2.

Аналогичная взаимосвязь между увеличением влажности и уменьшением твердости плодов авокадо выявлена в работах [115,127,149]. Калибровочные модели по исходным гиперспектральным изображениям могут быть использованы для определения влажности и сухого вещества в интервалах допустимых значений в соответствии с действующим стандартом *UNECE STANDARD FFV-42:2019* [17].

5.4 Применение вегетационных индексов гиперспектральных изображений для сортировки плодов авокадо

При классификации плодоовощной продукции по *HSI* изображению реализуют, в основном, два подхода: используют общее число каналов спектра или проводят выбор двух или более спектральных каналов, которые служат основой для расчета вегетационных индексов. Применением вегетационных индексов и *HSI* возможно определять порчу клубники [109] или цитрусовых [29].

Вегетационные индексы рассчитываются компьютерной программой, которая представляет *RGB*-изображение или изображение с использованием псевдоцвета.

В данном разделе исследования показана разработка методического подхода цифровой сортировки плодов авокадо по *HSI* с использованием вегетационных индексов.

Использован набор образцов авокадо, описанный в таблице 3.

Начиная с 7 дня хранения по 12 день наблюдался прогрессивный характер возникновения дефектов: значительное потемнение кожицы, возникновение на отдельных плодах грибковых заболеваний, а также значительное снижение плотности мезокарпия и возникновение внешней влажности [16].

Регистрация гиперспектральных изображений образцов плодов авокадо осуществлена следующим образом: для 7 свежих образцов в момент их получения, для 40 образцов, хранившихся в темном месте при комнатной температуре - на 7,9 и 12 дни хранения.

Для построения матрицы данных проведен отбор спектральных сигнатур гиперспектральных изображений поверхности плодов авокадо. При отборе спектральных сигнатур было использовано четыре гиперспектральных изображения плодов авокадо различного срока хранения: свежие образцы – *fresh*, образцы на 7,9,12 день хранения – *7d*, *9d*, *12d* соответственно. Отбор областей интереса как совокупности пикселей проведен по внутреннему контуру поверхности плодов авокадо. По отобраным *ROI* вычислены средние коэффициенты отражения на длинах волн спектра для каждого плода авокадо и осуществлен их импорт серии в формате *.csv* для последующего построения матрицы данных.

По результатам отбора построена матрица данных X размерностью $I \times J$ (I – количество *ROI* поверхности плодов авокадо различного срока хранения, J – число средних значений коэффициентов отражения на длинах волн (число каналов)). Обработка матрицы данных X методом главных компонент проведена в программном комплексе *Unscrambler X 10.0.4* (Camo Software, Норвегия). Визуализация *HSI* плодов авокадо по вегетационным индексам проведена в программном комплексе Альбеда 4.0.23 (МФТИ, Россия) [2].

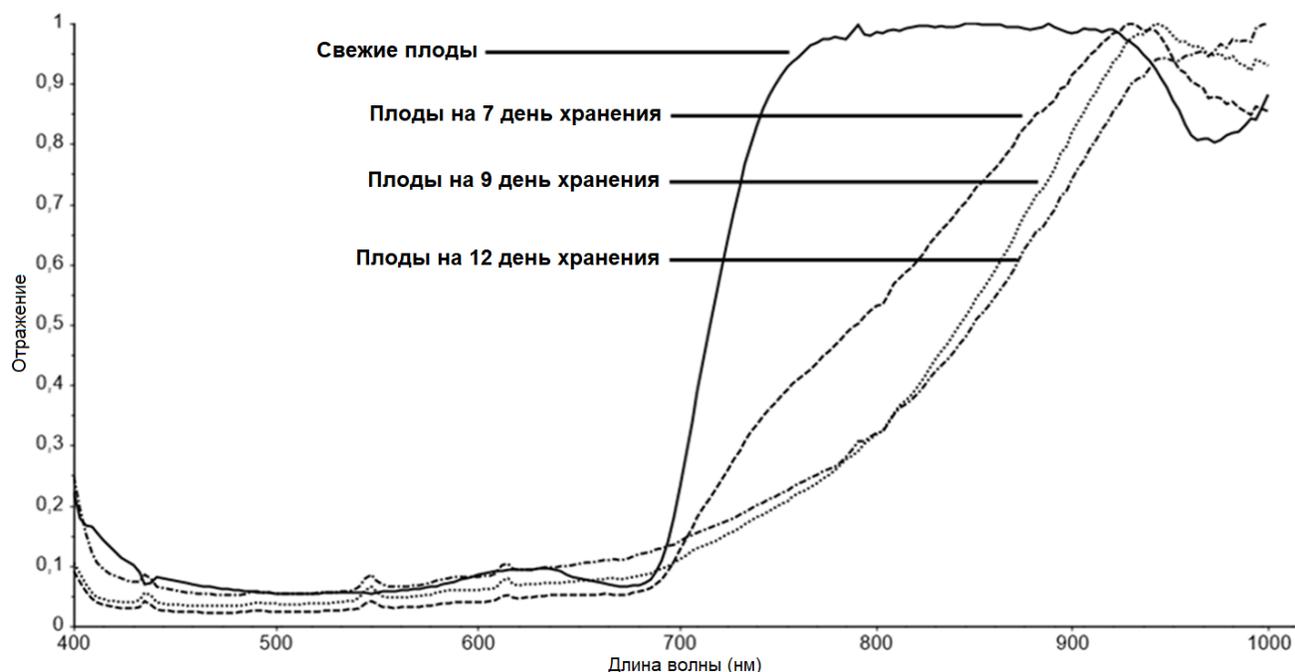


Рисунок 38 – Средние значения спектральных сигнатур четырех категорий плодов авокадо различного срока хранения: свежие плоды, плоды на 7,9,12 день хранения. Источник: составлено автором

Анализ спектральных сигнатур HSI плодов авокадо. На рисунке 38 представлено изменение усредненных спектральных сигнатур плодов авокадо от срока хранения. Как видно из рис. 38, в диапазоне 400-620нм видимой области их спектры слабо различимы, что подтверждается органолептической оценкой окраски плодов (ниже порога визуального различения). Но при этом, спектры значительно различаются в диапазоне «красного края» (*red-edge*) 680-780нм и в ближней инфракрасной области.

Для выявления наиболее значимых полос отражения в спектрах плодов авокадо проведена декомпозиция матрицы данных X методом главных компонент (МГК). На рис. 39 показаны значения факторных нагрузок длин волн спектра в координатах 1 и 2 главных компонент. Согласно знаку (\pm) и значению факторных нагрузок длин волн спектра проведена интерпретация ГК следующим образом:

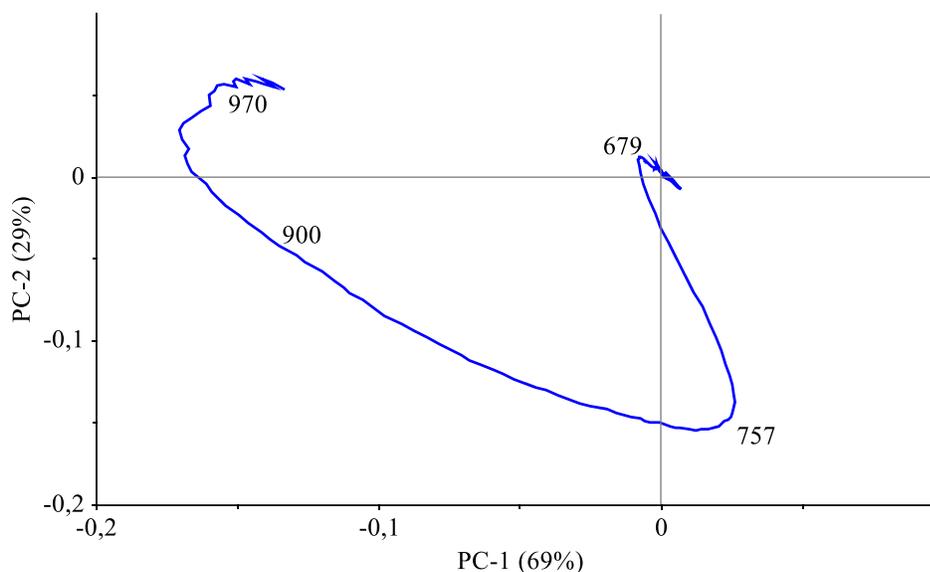


Рисунок 39 – Факторные нагрузки коэффициентов отражения спектра в диапазоне 400-1000 нм плодов авокадо различного срока хранения в координатах 1 и 2 ГК: *PC-1* – 1 ГК, *PC-2* – 2 ГК.

Источник: составлено автором

-основной вклад (69% от общей дисперсии) вносит 1 ГК, которая имеет максимальные значения факторных нагрузок длин волн со знаком (+) при 757 нм и со знаком (-) при 927 нм, показывающая изменения в ближней инфракрасной области спектра. Максимум при 927 нм, расположенный в области 1-го обертона колебаний ОН (900-970нм) [29,109], соответствует изменению содержания влаги в верхних слоях плодов авокадо при хранении;

-менее значим вклад 2 ГК (29%), которая имеет максимальные значения со знаком (\pm) при 679 и 757 нм [52], соответствующих диапазону «красного края» спектра. Выявленные по 2 ГК максимумы факторных нагрузок коэффициентов отражения длин волн спектра связаны с изменением и деградацией хлорофилла в экзокарпии при хранении плодов авокадо.

Основная область видимого спектра от 400 до 676 нм расположена на пересечении осей 1 и 2 ГК и, соответственно, вносит малый вклад в различие плодов авокадо от срока хранения, что упоминалось выше.

Максимум факторных нагрузок коэффициентов отражения спектра при 757 нм был использован для модификации вегетационного индекса *NDVI*, как показано ниже.

Вегетационные индексы. Методом числового интегрирования был произведён расчет индекса *CFI*. Индекс *CFI* (*Consumer Freshness Index*) или Индекс потребительской свежести, который характеризует степень пригодности для употребления человеком плодов авокадо. Чем выше значение индекса, тем выше потребительская свежесть плода и наоборот. Индекс *CFI* был рассчитан с помощью метода числового интегрирования

$$CFI = \int_i^n R_i, \quad (4)$$

где $i=680$, $n=800$, R_i – значение коэффициента отражения при соответствующей длине волны спектра i .

Один из наиболее широко используемых индексов, вычисляемый из многоспектральной информации, считается индекс нормализованного различия произрастания или *Normalized Difference Vegetation Index (NDVI)*. *NDVI* вычисляется, как:

$$NDVI = \frac{R_{800} - R_{680}}{R_{800} + R_{680}}, \quad (5)$$

где R_{800} , R_{680} – спектральное отражение на 800 нм и на 680 нм, соответственно [9,74].

На основе анализа факторных нагрузок длин волн спектра проведена модернизация индекса *NDVI*. В отличие от стандартного индекса *NDVI*, в котором точкой экстремума диапазона *red-edge* является полоса при 800 нм, в предлагаемом индексе *NDVI757/679* за данную точку была отобрана полоса при 757 нм, которая обладает максимальной дисперсией и, соответственно, сильнее связана с различиями в спектральных сигнатурах образцов.

$$NDVI757/679 = \frac{R_{757} - R_{679}}{R_{757} + R_{679}}, \quad (6)$$

где R_{757} и R_{679} – это спектральное отражение на 757 нм и на 679 нм соответственно.

Индекс вегетации растительной воды или *Plant Water Index (PWI)* количественно показывает содержание воды в исследуемых образцах и рассчитывается по формуле:

$$PWI = \frac{R_{970}}{R_{900}}, \quad (7)$$

где R_{970} и R_{900} – спектральное отражение на 970нм и на 900нм соответственно [81].

Расчет вегетационных индексов и оценка их эффективности. По данным матрицы данных X спектральных сигнатур плодов авокадо рассчитаны вегетационные индексы. В табл. 17 представлены средние значения индексов для различных категорий образцов авокадо.

При увеличении срока хранения образцов значение вегетационных индексов $NDVI$ и $NDVI_{757/679}$ уменьшаются, а значение индекса растительной воды PWI – увеличивается (рис. 40).

Таблица 17 – Средние значения вегетационных индексов образцов авокадо от срока хранения

Индекс	Категории			
	<i>fresh</i>	<i>7d</i>	<i>9d</i>	<i>12d</i>
<i>CFI</i>	35,570	26,41	19,47	18,23
<i>NDVI</i>	0,858	0,74	0,66	0,52
<i>NDVI_{757/679}</i>	0,856	0,66	0,50	0,36
<i>PWI</i>	0,850	0,96	1,06	1,12

Источник: составлено автором

При увеличении значений индекса PWI уменьшаются значения индексов $NDVI$ и $NDVI_{757/679}$. PWI коррелирует с изменением влаги у плодов авокадо, а $NDVI$, в свою очередь, связан с изменением пигментов при созревании и хранении плодов авокадо. При длительном хранении авокадо происходит его порча с одновременным

накоплением влаги в верхних слоях плодов и изменением окраски – деградацией хлорофилла. Таким образом, индексы *PWI*, *NDVI* и *NDVI757/679* описывают характерные изменения физиологического состояния плодов авокадо при длительном хранении. При этом выявлено, что индекс *NDVI757/679* обладает более высокой обратной взаимосвязью с *PWI* ($R^2=0,99$), чем *NDVI* с *PWI* ($R^2=0,95$) [16].

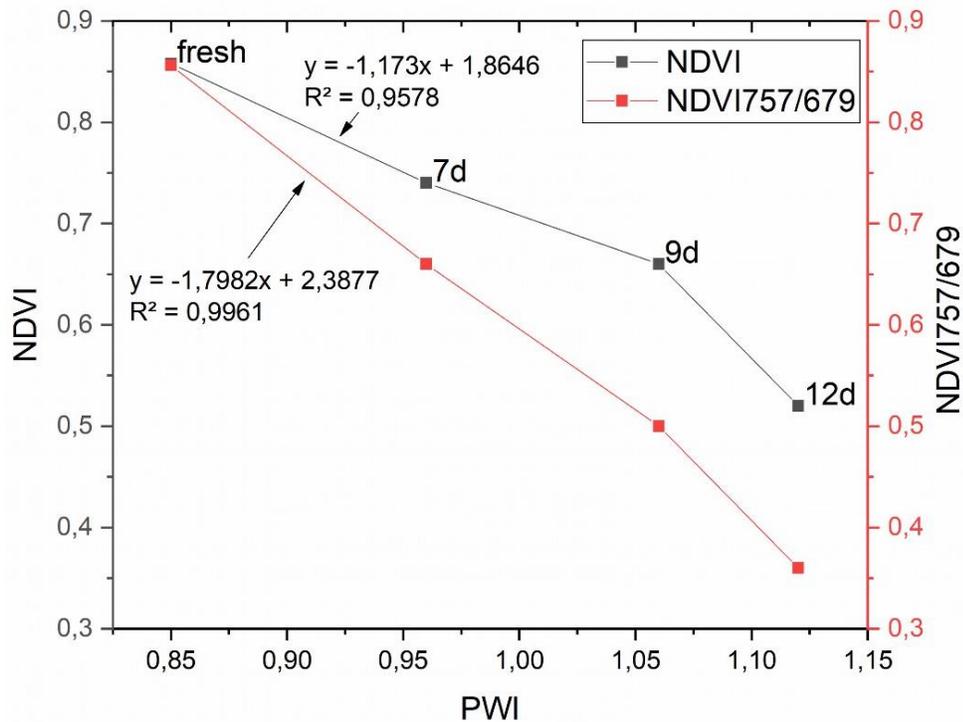


Рисунок 40 – Изменение соотношения значения индексов *NDVI*, *NDVI757/679* и *PWI* от срока хранения плодов авокадо на 7, 9 и 12 день хранения

Источник: составлено автором

Визуализация *HSI* плодов авокадо с помощью вегетационных индексов. Результатом применения двух вегетационных индексов является визуализация *HSI* изображений состояния поверхности плодов авокадо [16].

Визуализация гиперспектральных изображений плодов авокадо различного срока хранения, построенные с помощью двух вегетационных индексов: *PWI* и усо-

вершенствованного $NDVI_{757/679}$, в оттенках серого цвета (от темно-серого с минимальными значениями индексов до светло-серого с максимальными значениями индексов) представлены на рис. 41 и 42.

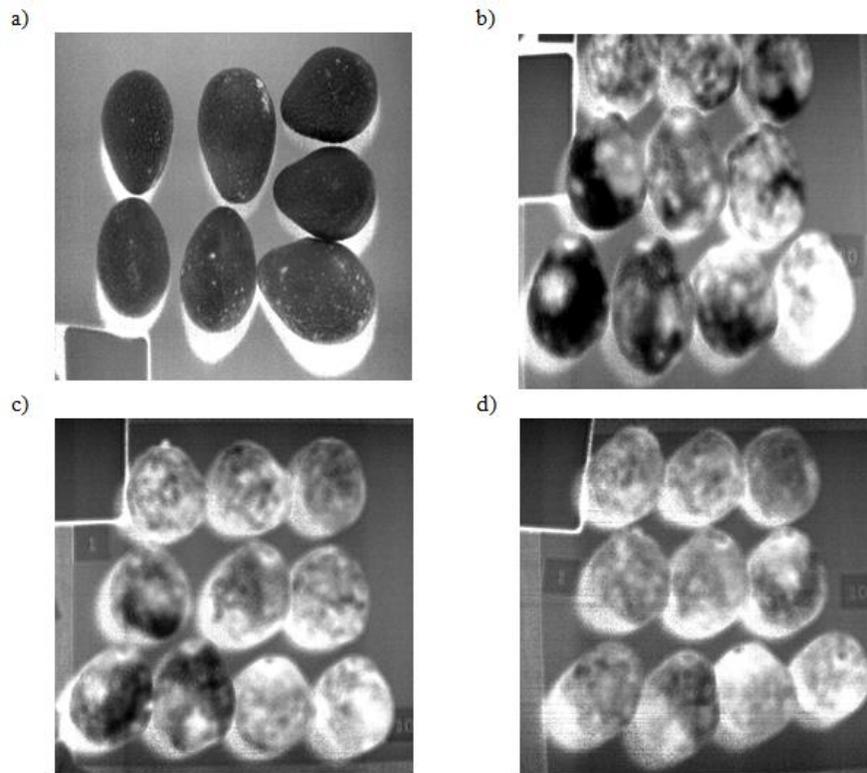


Рисунок 41 – Гиперспектральные изображения плодов авокадо после применения индекса PWI : образцов-*fresh* (a); образцов на 7 день хранения (b); образцов на 9 день хранения (c); образцов на 12 день хранения (d)

Источник: составлено автором

Повышение площади светлых участков на поверхности плодов с увеличением срока хранения (рис. 41) соответствует увеличению содержания влаги в верхних слоях плодов (таблица 17, рис. 40).

Наоборот, с повышением площади темно-серых участков на поверхности плодов при увеличении срока хранения (рис. 42) также соответствует увеличению содержания влаги в верхних слоях плодов (таблица 17, рис. 40). Установлена обратная за-

висимость между двумя вегетационными индексами PWI и $NDVI_{757/679}$ (рис. 40). Сопоставляя результаты анализа данных, представленных рис. 41 и 42, можно отметить, что гиперспектральные изображения, построенные по вегетационному индексу PWI обратны по псевдоцвету в оттенках серого индексу $NDVI_{757/679}$. Эти обратно связанные вегетационные индексы хорошо отражают состояние плодов авокадо от срока их хранения, что обусловлено увеличением влажности [16].

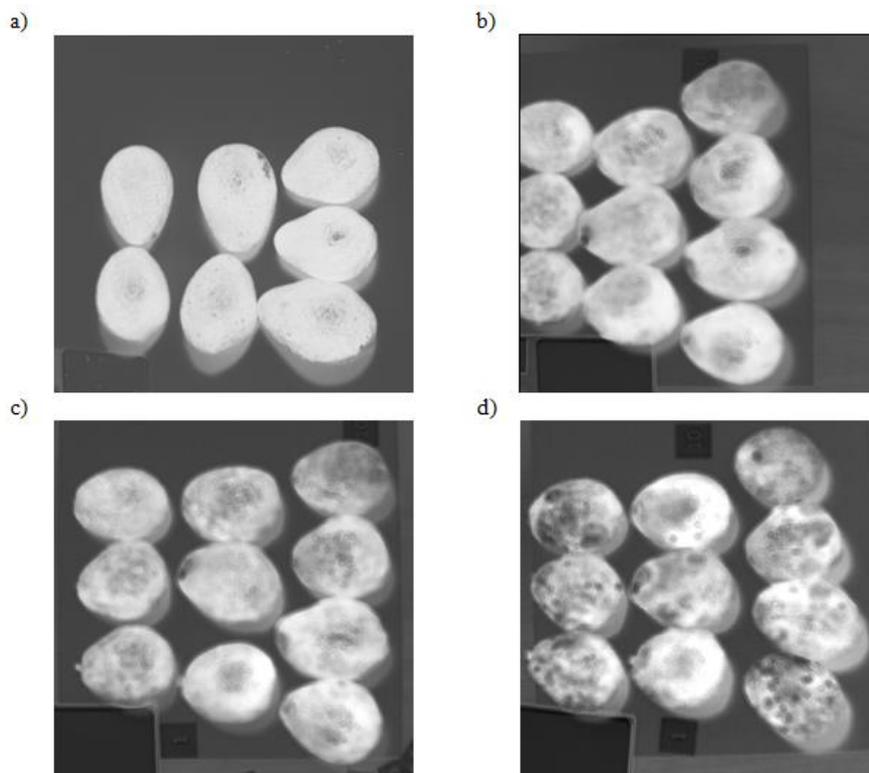


Рисунок 42 – Гиперспектральные изображения плодов авокадо после применения индекса $NDVI_{757/679}$: образцов-*fresh* (a); образцов на 7 день хранения (b); образцов на 9 день хранения (c); образцов на 12 день хранения (d)

Источник: составлено автором

Таким образом, использование гиперспектральных камер позволяет определять диапазон потребительской зрелости плодов авокадо в период их реализации в коммерческих организациях в режиме реального времени на месте без их разрушения [16].

По результатам исследования разработан методический подход анализа гиперспектральных изображений, полученных использованием портативной (400-1000 нм) и лабораторной (935-1720 нм) камер, для контроля качества пищевой продукции:

-для сортировки зерен гречихи по степени выполненности на основании отбора спектральных сигнатур из гиперспектрального изображения в диапазоне 935-1720 нм разработана классификационная модель PLS-DA.

-для обнаружения визуально не выявляемых дефектов авокадо разработаны компьютерные модели классификации плодов авокадо по наличию дефектов методами *PLS-DA*, *SIMCA*;

-для определения влажности плодов авокадо посредством отбора спектральных сигнатур *HSI* случайным образом со всей поверхности плодов или всей поверхности изображения *HSI* плодов как области интереса и компьютерного моделирования;

-для сортировки плодов авокадо по степени зрелости путем расчета и модификации вегетационных индексов (*PWI* и *NDVI757/679* соответственно) и визуализации гиперспектрального изображения по их значениям;

Заключение

1) Разработаны и научно обоснованы методические подходы применения неразрушающих методов оптической спектроскопии, гиперспектральных изображений и многомерного анализа для разработки компьютерных моделей идентификации и контроля качества пищевой продукции (зерно гречихи, гречневая крупа, плоды авокадо, зерновой кофе, сливочное масло).

2) Сформированы базы данных: *UV-VIS-NIR*-спектров (350-2500 нм), ИК-спектров ($4000-400\text{ см}^{-1}$) и цветовых характеристик в системе CIE $L^*a^*b^*$, гиперспектральных изображений в видимом (400-1000 нм) и ближнем инфракрасном диапазоне (935-1720 нм) отдельных видов пищевых продуктов. Анализ *2D*-матриц спектроскопии и *3D*-матриц гиперспектральных изображений проведен методами многомерной классификации (МГК, ДА, *SIMCA*, *PLS-DA*, *PCA-DA*) и градуировки (РГК, *PLS*). При использовании гиперспектральных изображений отработан следующий методический подход: отбор спектральных сигнатур *ROI*, сегментация изображений, визуализация гиперспектральных изображений с использованием в качестве псевдоцветов длин волн и главных компонент.

3) Продемонстрировано применение оптической спектроскопии и многомерного анализа для разработки моделей классификации пищевых продуктов на категории качества: 1) зернового кофе по ботаническому виду, обжарке и географическому региону произрастания (Америка, Азия, Африка); 2) гречневой крупы по времени урожая (весенний, осенний) и в зависимости от гидротермической обработки.

4) Отработан методический подход применения гиперспектральных изображений в видимой области для разработки компьютерных моделей идентификации пищевой продукции: 1) Анализ дисперсии данных гиперспектрального изображения плодов авокадо Хасс позволил выявить полосы и диапазоны спектра в видимой области,

объясняющие физиологическое состояние плодов и являющиеся критериями идентификации. Методом *SIMCA* разработана модель классификации плодов авокадо Хасс на категории качества по наличию дефектов. Разработаны и верифицированы калибровочные модели определения влажности и сухого вещества в составе плодов авокадо по исходным гиперспектральным изображениям. Продемонстрирован методический подход цифровой сортировки авокадо по сроку хранения с использованием вегетационных индексов и визуализации гиперспектральных изображений по их значениям; 2) Проведен анализ гиперспектральных изображений в ближнем инфракрасном диапазоне зерна гречихи. На основании рентгенообразов зерна гречихи образцы были распределены на группы по степени выполненности. С помощью отбора отдельных спектров из гиперспектрального изображения и расчета методом *PLS-DA* разработана многомерная классификационная модель градации зерна гречихи по степени выполненности.

5) Разработаны и апробированы калибровочные модели прогнозирования свойств пищевых продуктов: 1) модель для выявления ассортиментной фальсификации сливочного масла пальмовым; 2) модель для прогнозирования кислотного числа жира и сроков хранения гречневой крупы.

б) Результаты работы апробированы в лаборатории качества *Q-lab* сети «Азбука Вкуса» при проведении входного контроля качества партии авокадо, при проведении лекционных и практических занятий, подготовке курсовых и выпускных квалификационных работ бакалавров и магистров в рамках учебного процесса по направлениям подготовки «Товарный менеджмент и экспертиза», «Пищевые инновации и продакт-дизайн» ФГБОУ ВО «РЭУ им. Г.В. Плеханова».

Список сокращений и условных обозначений

ИК	– инфракрасный
БИК	– ближний инфракрасный (<i>NIR</i>)
ГК	– главная компонента
ДА	– дискриминантный анализ
КА	– кластерный анализ
МГК	– метод главных компонент
ПЛС	– проекция на латентные структуры (<i>PLS</i>)
РГК	– регрессия на главные компоненты
УФ	– Ультрафиолетовый (<i>UV</i>)
<i>CIEL*a*b*</i>	– колориметрическая система (<i>L*</i> -ось светлота; ось <i>a*</i> - <i>+a*</i> -красный, <i>-a*</i> - зеленый; ось <i>b*</i> - <i>+b*</i> -желтый, <i>-b*</i> синий)
<i>DM, %</i>	– сухой остаток (<i>dry matter</i>)
<i>HSI</i>	– гиперспектральное изображение (<i>hyperspectral imaging</i>)
<i>MS, %</i>	– влажность, (<i>moisture</i>)
<i>MSC</i>	– мультипликативная коррекция рассеяния света (<i>multiplicative scatter correction</i>)
<i>PCA-LDA</i>	– метод главных компонент-линейный дискриминантный анализ (<i>principal component analysis-linear discriminant analysis</i>)
<i>PCA-QDA</i>	– метод главных компонент-квадратичный дискриминантный анализ (<i>principal component analysis-quadratic discriminant analysis</i>)
<i>PLS-DA</i>	– дискриминантный анализ при помощи проекции на латентные структуры (<i>partial least squares discriminant analysis</i>)
<i>SIMCA</i>	– формальное независимое моделирование аналогий классов (<i>soft independent modelling of class analogies</i>)

- SNV** – коррекция стандартная нормальная вариация (*Standard Normal Variate*)
- SVR** – опорный вектор регрессии (*support vector regression*)
- R^2** – коэффициент детерминации. R_c^2 / R_v^2 - коэффициент детерминации калибровки/ валидации
- Red edge** – диапазон «красный край» (690-780 нм)
- RMSE** – среднеквадратичная ошибка (*Root Mean Square Error*). *RMSEC*/*RMSEV* - среднеквадратичная ошибка калибровки/ валидации
- ROI** – область интереса (*region of interest*)
- VIS** – видимая область

Список использованной литературы

1. Авилова, И. А. Возможность использования метода ИК-спектроскопии при анализе сырья и продуктов растительного происхождения / И. А. Авилова, Д. В. Хлыстов // Известия Юго-Западного государственного университета. Серия: Физика и химия. – 2014. – № 1. – С. 34-38. – ISSN 2223-1544. – Текст : электронный.
2. Альбедро - программное обеспечение для обработки гиперспектральных данных. Кафедра СУМГФ МФТИ. — URL : <https://geo.mipt.ru/albedo> (дата обращения: 17.02.2021). – Текст : электронный.
3. Белецкий, С. Л. Новый аппаратно-программный рентгенодиагностический комплекс / С. Л. Белецкий, Н. А. Хаба, О. С. Шилкова // Инновационные технологии производства и хранения. – 2018. – 9(9). – С. 26-32. – ISBN 978-5-906936-24-0. – Текст : электронный.
4. Буданина, Л. Н. Определение состава спредов сливочного масла с пальмовым методами термического анализа / Л. Н. Буданина, А. Л. Верещагин, Н. В. Бычин // Техника и технология пищевых производств. – 2015. – №. 3 (38). – С. 133-138. – Online ISSN: 2313-1748, Print ISSN: 2074-9414. – Текст : электронный.
5. Вытовтов, А. А. Определение подлинности и обнаружение фальсификации пищевых продуктов методом ИК-Фурье-спектрометрии / А. А. Вытовтов // Ученые записки СПб-филиала РТА. – 2010. – № 1 (35) – С. 193-196. – ISSN: 1684-1026. – Текст : электронный.
6. Дрейпер, Н., Смит, Г. Прикладной регрессионный анализ / 3-е изд. - Вильямс. - 2016 - 912 с. – ISBN 978-5-8459-0963-3. – Текст : непосредственный.
7. Золотов, Ю.А. Аналитическая химия: проблемы и достижения / Ю.А. Золотов // Наука, Москва. – 1992. – 284 с. – ISBN 5-020015-28-8. – Текст : непосредственный.

8. Исследование качества – масло сливочное 82,5 % жирности // Роскачество: сайт. 26.10.2016. – URL: <https://rskrf.ru/ratings/produkty-pitaniya/molochnye-produkty/maslo-slivochnoe-82-5-zhirnosti/> (дата обращения: 08.10.2018). – Текст : электронный.
9. Костин, И. Г. Возможности использования современных геоинформационных систем для агроэкологического мониторинга земель сельскохозяйственного назначения / И. Г. Костин // Достижения науки и техники АПК. – 2020. – Т. 34. – №. 9. – С. 96-105. – ISSN 0235-2451. – Текст : электронный.
10. Лобанова, А. А. Исследование биологически активных флавоноидов в экстрактах из растительного сырья / А.А. Лобанова, В.В. Будаева, Г.В. Сакович // Химия растительного сырья. – 2004. – № 1. – С. 47–52. – ISSN (Print): 1029-5151, ISSN (Online): 1029-5143. – Текст : электронный.
11. Монахова, Ю. Б. Методология хемометрического моделирования спектрометрических сигналов в анализе объектов сложного состава / Ю. Б. Монахова, С. П. Муштакова // Журнал аналитической химии. – 2017. – Т. 72, № 2. – С. 119-128. – ISSN 0044-4502 – Текст : непосредственный.
12. Муллина Э. Р., Чупрова Л. В. Современные физико-химические Методы анализа продуктов питания. – URL: <http://vernadsky.tstu.ru/pdf/2007/03/34.pdf> (дата обращения 23.09. 2019). – 2007. – Текст : электронный
13. Марьин, В. А. Динамика изменений механических свойств пропаренных зёрен гречихи, перезимовавшей на корню / В. А. Марьин, А. Л. Верещагин, Н. В. Бычин // Хлебопродукты. – 2019. – № 5. – С. 58-61. – ISSN 0235-2508. – Текст : электронный.
14. Марьин, В. А. Зерно и крупа весеннего урожая / В. А. Марьин, А. Л. Верещагин // Технологии и оборудование химической, биотехнологической и пищевой промышленности. – 2018. – С. 515-519. – ISBN 978-5-9257-0317-5. – Текст : электронный.

15. Марьин, В. А. Распределение размера ядра во фракциях зерна гречихи / В. А. Марьин, А. Л. Верещагин // Хранение и переработка сельхозсырья. – 2019. – № 1. – С. 130-138. – ISSN 2072-9669 (Print), ISSN 2658-767X (Online). – Текст : электронный.
16. Метленкин, Д.А. Применение вегетационных индексов гиперспектральных изображений для сортировки плодов авокадо / Д.А. Метленкин [и др.] // Достижения науки и техники АПК. – 2023. – Т. 37. – № 1. – С. 41–45. – ISSN 0235-2451. – Текст : электронный.
17. Метленкин Д.А. Сортировка плодов авокадо по гиперспектральным изображениям / Д.А. Метленкин [и др.] // Пищевые системы. – 2023. – Т. 6. – № 1. – С. 46-53. – ISSN 2618-9771. – Текст : электронный.
18. Метленкин, Д.А. Хемометрическое моделирование спектрометрических данных для идентификации и прогнозирования свойств продовольственных продуктов / Д.А. Метленкин, Ю.Т. Платов, Р.А. Платова // Региональные рынки потребительских товаров: качество, экологичность, ответственность бизнеса. Материалы III Всероссийской научно-практической конференции с международным участием. Новосибирск, 2020. – 2021. – С. 91-95. – ISBN 978-5-7638-4461-0. – Текст : электронный.
19. Нечипоренко, А. П. Исследование растительных масел и их купажей методами инфракрасной спектроскопии отражения и рефрактометрии / А. П. Нечипоренко [и др.] // Научный журнал НИУ ИТМО. Серия «Процессы и аппараты пищевых производств». – №1 – 2018. – С. 3-14. – ISSN 2310-1164. – Текст : электронный.
20. Новые физико-химические и биотехнологические методы обработки пищевого сырья и продуктов: учебное пособие для обучающихся по программе магистратуры 19.04.03 Продукты питания животного происхождения / сост.: А.Л. Алексеев; Донской ГАУ. – Персиановский : Донской ГАУ, 2019 – 183 с. – Тираж 300 экз. – Текст : непосредственный.

21. ООО "Компания "АЗИМУТ ФОТОНИКС" [Электронный ресурс] / Гиперспектральные камеры. – URL: https://azimp.ru/catalogue/research_and_industry/ (Дата обращения: 15.03.2023).
22. Пенджиев, А. М. Физико-химический состав кофейного напитка / А. М. Пенджиев // Рациональное питание, пищевые добавки и биостимуляторы. – 2016. – № 1. – С. 100-111. – ISSN 2414-1054. – Текст : электронный.
23. Платов, Ю. Т. Идентификация гречневой крупы сочетанием методов УФ-видимой-ближней ИК-спектроскопии и многомерного анализа / Ю. Т. Платов [и др.] // Журнал прикладной спектроскопии. – 2021. – Т. 88. – № 4. – С. 541-549. – ISSN 0514-7506. – Текст : электронный.
24. Платов, Ю. Т. Спектроколориметрический метод идентификации и контроля качества гречневой крупы / Ю. Т. Платов, Д. А. Метленкин, Г. А. Бобожонова // Хлебопечение России. – 2020. – №1. – С. 24-28. – ISSN 2073-3569. – Текст : непосредственный.
25. Популярное введение в современный анализ данных и машинное обучение на STATISTICA / В.П. Боровиков. – М.: ИнфоКомкнига, 2018. – 354 с. – ISBN. 978-5-9912-0326-5. – Текст : электронный.
26. Родионова, О. Е. Хемометрика: достижения и перспективы / О. Е. Родионова, А. Л. Померанцев // Успехи химии. – 2006. – №. 75(4). – С. 302-321. – ISSN 0042-1308 – Текст : непосредственный.
27. Эсбенсен, К. Анализ многомерных данных / К. Эсбенсен // ИПХФ РАН, Черноголовка. – 2005. – 160 с. – ISBN 5-901675-50-9. – Текст : непосредственный.
28. Ягофаров, Д. Ш. Физико-химические свойства картофельного крахмала / Д. Ш. Ягофаров [и др.] // Вестник Казанского технологического университета. – 2012. – Т. 15. – №. 12. – С. 212-215. – ISSN 1998-7072. – Текст : электронный.
29. Abdulridha, J. UAV-based remote sensing technique to detect citrus canker disease utilizing hyperspectral imaging and machine learning / J. Abdulridha, O. Batuman, Y.

Ampatzidis // *Remote Sensing*. – 2019. – Vol. 11. – №. 11. – P. 1373. – ISSN 2072-4292. – Текст : электронный.

30. Ahmed, A. Phytochemicals and biofunctional properties of buckwheat: a review / A. Ahmed [et al.] // *The Journal of Agricultural Science*. – 2014. – Vol. 152. – №. 3. – P. 349-369. – ISSN (Print): 0021-8596, ISSN (Online): 1469-5146. – Текст : электронный.

31. Arslan, F. N. Rapid detection of authenticity and adulteration of cold pressed black cumin seed oil: A comparative study of ATR-FTIR spectroscopy and synchronous fluorescence with multivariate data analysis / F. N. Arslan [et al.] // *Food control*. – 2019. – 98. – P. 323-332. – ISSN 0956-7135. – Текст : электронный.

32. Ashton, O. B. Pigments in avocado tissue and oil / O. B. Ashton [et al.] // *Journal of Agricultural and Food Chemistry*. – 2006. – Vol. 54. – №. 26. – P. 10151-10158. – Online ISSN: 1520-5118, Print ISSN: 0021-8561. – Текст : электронный.

33. AT, D. Scientific concepts of functional foods in Europe consensus document / D. AT, E. C. Action // *British Journal of Nutrition*. – 1999. – Vol. 81. – P. S1-S27. – ISSN (Print): 0007-1145, ISSN (Online): 1475-2662. – Текст : электронный.

34. Bannon, D. Cubes and slices: Imaging / D. Bannon // *Nature photonics (Print)*. – 2009. – Vol. 3. – №. 11. – P. 627-629. – ISSN 1749-4885. – Текст : непосредственный.

35. Barron, C. Prediction of relative tissue proportions in wheat mill streams by fourier transform mid-infrared spectroscopy / C. Barron // *Journal of Agricultural and Food Chemistry*. – 2011. – Vol. 59(19). – P. 10442-10447. – Online ISSN: 1520-5118, Print ISSN: 0021-8561. – Текст : электронный.

36. Behmann, J. Specim IQ: evaluation of a new, miniaturized handheld hyperspectral camera and its application for plant phenotyping and disease detection / J. Behmann [et al.] // *Sensors*. – 2018. – Vol. 18. – №. 2. – P. 441. – Online ISSN: 1424-8220. – Текст : электронный.

37. Bellincontro, A. Feasible application of a portable NIR-AOTF tool for on-field prediction of phenolic compounds during the ripening of olives for oil production / A. Bellincontro [et al.] // *Journal of Agricultural and Food Chemistry*. – 2012. – 60(10). – P. 2665-2673. – Online ISSN: 1520-5118, Print ISSN: 0021-8561. – Текст : электронный.
38. BeMiller, J. *Starch: Chemistry and Technology* / J. BeMiller // 3rd edition. – Burlington, USA: Academic Press. – 2009. – 894 p. – ISBN 978-0-12-746275-2. – Текст : непосредственный.
39. Bhupathiraju, S. N. Caffeinated and caffeine-free beverages and risk of type 2 diabetes / S. N. Bhupathiraju [et al.] // *The American journal of clinical nutrition*. – 2013. – Vol. 97. – №. 1. – P. 155-166. – ISSN 0002-9165 (Print), ISSN 1938-3207 (Online). – Текст : электронный.
40. Błaszczak, W. Antioxidant Properties and Rutin Content of High Pressure-Treated Raw and Roasted Buckwheat Groats / W. Błaszczak [et al.] // *Food and Bioprocess Technology*. – 2011. – № 6. – P. 92–100. – ISSN 1935-5130. – Текст : электронный.
41. Bolton, J. F. Full spectral imaging: a revisited approach to remote sensing / J. F. Bolton // *Sensors, Systems, and Next-Generation Satellites VII*. – SPIE, 2004. – Vol. 5234. – P. 243-251. – ISSN 1996-756X. – Текст : электронный.
42. Briandet, R. Approaches to adulteration detection in instant coffees using infrared spectroscopy and chemometrics / R. Briandet, E. K. Kemsley, R. H. Wilson // *Journal of the Science of Food and Agriculture*. – 1996. – Vol. 71. – P. 359-366. – ISSN 2575-7954. – Текст : электронный.
43. Cano-Marquina, A. The impact of coffee on health / A. Cano-Marquina, J. J. Tarín, A. Cano // *Maturitas*. – 2013. – Vol. 75. – №. 1. – P. 7-21. – ISSN 0378-5122 (Print), ISSN 1873-4111 (Online). Текст : электронный.
44. Cayuela, J. A. Rapid NIR determination of alkyl esters in virgin olive oil / J. Cayuela // *Grasas Y Aceites*. – 2017. – 68(2). – P. 195. – ISSN 0017-3495. – Текст : электронный.

45. Cen, H. Theory and application of near infrared reflectance spectroscopy in determination of food quality / H. Cen, Y. He // *Trends in Food Science & Technology*. – 2007. – Vol. 18. – №. 2. – P. 72-83. – ISSN 0924-2244. – Текст : электронный.
46. Cerretani, L. Rapid FTIR determination of water, phenolics and antioxidant activity of olive oil / L. Cerretani [et al.] // *European Journal of Lipid Science and Technology*. – 2010. – 112(10). – P. 1150-1157. – ISSN: 1438-9312 (online). – Текст : непосредственный.
47. Clark, C. J. Dry matter determination in ‘Hass’ avocado by NIR spectroscopy / C. J. Clark [et al.] // *Postharvest Biology and Technology*. – 2003. – Vol. 29. – №. 3. – P. 301-308. – Print ISSN: 0925-5214, Online ISSN: 1873-2356. – Текст : электронный.
48. Clément, A. Nondestructive Measurement of Fresh Tomato Lycopene Content and Other Physicochemical Characteristics Using Visible-NIR Spectroscopy / A. Clément, M. Dorais, M. Vernon // *Journal of Agricultural and Food Chemistry*. – 2008. – 56. – P. 9813–9818. – Online ISSN: 1520-5118, Print ISSN: 0021-8561. – Текст : электронный.
49. Colour in food. Improving quality / Ed. D.B. MacDougall. CRC Press, Cambridge, UK. –2002. – 378 p. – ISBN 978-1-8557-3590-3. – Текст : электронный.
50. Corpaş, L. Karl Fischer Water Titration– Principal Component Analysis Approach on Wheat Flour / L. Corpaş [et al.] // *Food Analytical Methods*. – 2014. – Vol. 7. – №. 6. – P. 1353-1358. – ISSN 1936-9751. – Текст : непосредственный.
51. Correa, E. C. Optimal management of oil content variability in olive mill batches by NIR spectroscopy / E. C. Correa [et al.] // *Scientific Reports*. – 2019. – 9. – article 13974. – Online ISSN: 2045-2322. – Текст : электронный.
52. Cox, K. A. Skin colour and pigment changes during ripening of ‘Hass’ avocado fruit / K. A. Cox [et al.] // *Postharvest Biology and Technology*. – 2004. – Vol. 31. – №. 3. – P. 287-294. – Print ISSN: 0925-5214, Online ISSN: 1873-2356. – Текст : электронный.
53. Craig, A. P. Discrimination between defective and non-defective roasted coffees by diffuse reflectance infrared Fourier transform spectroscopy / A. P. Craig, A. S.

Franca, L. S. Oliveira // LWT—Food Science and Technology. – 2012. – Vol. 47. – P. 505-511. – ISSN 0023-6438. Текст : электронный.

54. Cremer, D. R. Fourier transform infrared microspectroscopic study of the chemical microstructure of corn and oat flour-based extrudates / D. R. Cremer, G. Kaletunç // Carbohydrate Polymers. – 2003. – Vol. 52. – P. 53-65. – ISSN 0144-8617. – Текст : электронный.

55. Croft, H. Leaf Pigment Content / H. Croft, J. Chen // Comprehensive Remote Sensing. – 2017. – V.3. – P. 117–142. – ISBN 978-0-12-803221-3. – Текст : электронный.

56. Cubero, S. Advances in machine vision applications for automatic inspection and quality evaluation of fruits and vegetables / S. Cubero [et al.] // Food and Bioprocess Technology. – 2011. – 4. – P. 487-504. – ISSN 1935-5130. – Текст : непосредственный.

57. Cuibus, L. Preliminary Discrimination of Butter Adulteration by ATR-FTIR Spectroscopy / L. Cuibus [et al.] // Bulletin of University of Agricultural Sciences and Veterinary Medicine Cluj-Napoca Food Science and Technology. – 2015. – 72 (1). – P. 70-76. – ISSN 2344-2344. – Текст : электронный.

58. Curran, P. J. Exploring the relationship between reflectance red edge and chlorophyll content in slash pine / P. J. Curran, J. L. Dungan, H. L. Gholz // Tree physiology. – 1990. – Vol. 7. – №. 1-2-3-4. – P. 33-48. – ISSN (Print): 1568-2544, ISSN (Online) 1758-4469. – Текст : электронный.

59. Davey, M. W. Application of visible and near-infrared reflectance spectroscopy (Vis/NIRS) to determine carotenoid contents in banana (*musa spp.*) fruit pulp / M. W. Davey [et al.] // Journal of Agricultural and Food Chemistry. – 2009. – 57. – P. 1742-1751. – Online ISSN: 1520-5118, Print ISSN: 0021-8561. – Текст : электронный.

60. del Río F. J., Riu J., Rius F. X. Prediction intervals in linear regression taking into account errors on both axes // Journal of chemometrics. – 2001. – Vol. 15. – №. 10. – P. 773-788. – ISSN 0886-9383. – Текст : непосредственный.

61. Diviš, P. The effect of coffee beans roasting on its chemical composition / P. Diviš, J. Pořízka, J. Kříkala // *Potravinářstvo*. – 2019. – Vol. 13. – №. 1. – P. 344-350. – ISSN (Online) 1337-0960. – Текст : электронный.
62. Donetti, M. Biochemical markers defining growing area and ripening stage of imported avocado fruit cv. Hass / M. Donetti, L. A. Terry // *Journal of Food Composition and Analysis*. – 2014. – Vol. 34. – №. 1. – P. 90-98. – Print ISSN: 0889-1575, Online ISSN: 1096-0481. – Текст : электронный.
63. Dziedzic, K. Content of phytosterols in raw and roasted buckwheat groats and by-products / K. Dziedzic [et al.] // *Czech Journal of Food Sciences*. – 2015. – Vol. 33. – №. 5. – P. 424-430. – ISSN 1212-1800. – Текст : электронный.
64. Ertlen, D. Discriminating between organic matter in soil from grass and forest by near-infrared spectroscopy / D. Ertlen [et al.] // *European Journal of Soil Science*. – 2010. – Vol. 61. – №. 2. – P. 207-216. – ISSN 1351-0754. – Текст : электронный.
65. Esbensen, K. H. *Multivariate Data Analysis in Practice: An Introduction to Multivariate Data Analysis and Experimental Design* / K. H. Esbensen // 4 edition. – Oslo: Camo ASA. – 2000. – 352 p. – ISBN 978-8-2993-3303-0. – Текст : непосредственный.
66. Esteban-Díez, I. Coffee varietal differentiation based on near-infrared spectroscopy / I. Esteban-Díez [et al.] // *Talanta*. – 2007. – Vol. 71(1). – P. 221-229. – ISSN 0039-9140. – Текст : непосредственный.
67. Esteban-Díez, I. Generalization of OWAVEC method for simultaneous noise suppression, data compression and orthogonal signal / I. Esteban-Díez, J. M. González-Sáiz, C. Pizarro // *Analytica Chimica Acta*. – 2005. – Vol. 554 (1-2). – P. 89-99. – ISSN 0003-2670. – Текст : непосредственный.
68. Fearn, T. Assessing calibrations: SEP, RPD, RER and R2 / T. Fearn // *NIR News*. – 2002. – Vol. 13. – P. 12-14. – ISSN 0960-3360. – Текст : непосредственный.
69. Ferreira, D. S. Comparison and application of near-infrared (NIR) and mid-infrared (MIR) spectroscopy for determination of quality parameters in soybean samples / D.

S. Ferreira [et al.] // *Food Control*. – 2014. – Vol. 35. – №. 1. – P. 227-232. – ISSN 0956-7135. – Текст : непосредственный.

70. Ferreira, D. S. Fourier transform near-infrared spectroscopy (FT-NIRS) application to estimate Brazilian soybean [*Glycine max* (L.) Merrill] composition / D. S. Ferreira, J. A. Pallone, R. J. Poppi // *Food Research International*. – 2013. – Vol. 51. – №. 1. – P. 53-58. – ISSN 0963-9969. – Текст : непосредственный.

71. Ferreira da Vinha, A. Physicochemical parameters, phytochemical composition and antioxidant activity of the algarvian avocado (*Persea americana* Mill.) / A. Ferreira da Vinha, J. Moreira, S. Barreira // *The Journal of Agricultural Science*. – 2013. – Vol. 5. – №. 12. – P. 100-109. – ISSN (Print): 0021-8596, ISSN (Online) 1469-5146. – Текст : электронный.

72. Fu, X. Application of nir spectroscopy for firmness evaluation of peaches / X. Fu [et al.] // *Journal of Zhejiang University SCIENCE B*. – 2008. – 9. – P. 552-557. – ISSN1673-1581. – Текст : электронный.

73. Gallardo-Velazquez, T. Determination of ternary solutions concentration in liquid–liquid extraction by the use of attenuated total reflectance-Fourier transform infrared spectroscopy and multivariate data analysis / T. Gallardo-Velazquez [et al.] // *The Canadian Journal of Chemical Engineering*. – 2008. – 86(1). – P. 77-83. – ISSN 0008-4034. – Текст : электронный.

74. Gandhi, G. M. Ndvi: Vegetation change detection using remote sensing and gis—A case study of Vellore District / G. M. Gandhi [et al.] // *Procedia computer science*. – 2015. – Vol. 57. – P. 1199-1210. – ISSN 1877-0509. – Текст : электронный.

75. García Martín, J. F. Determination of the Acidity of Waste Cooking Oils by Near Infrared Spectroscopy / J. F. García Martín [et al.] // *Processes*. – 2019. – Vol. 7. – №. 5. – P. 304. – Online ISSN: 2227-9717. – Текст : электронный.

76. Garrigues, J. M. Fourier transform infrared determination of caffeine in roasted coffee samples / J. M. Garrigues [et al.] // *Fresenius Journal of Analytical Chemistry*. – 2000.

– Vol. 366. – P. 319-322. – ISSN 0937-0633 (Print), ISSN 1432-1130 (Online). – Текст : электронный.

77. Geladi, P. Linearization and scatter-correction for near-infrared reflectance spectra of meat / P. Geladi [et al.] // *Applied Spectroscopy*. – 1985. – №. 39(3). – P. 491-500. – ISSN 0003-7028. – Текст : непосредственный.

78. Geladi, P. Some recent trends in the calibration literature / P. Geladi // *Chemometrics and Intelligent Laboratory Systems*. – 2002. – №. 60. – P. 211-224. – ISSN 0169-7439. – Текст : непосредственный.

79. González-Caballero, V. First steps towards the development of a non-destructive technique for the quality control of wine grapes during on-vine ripening and on arrival at the winery / V. González-Caballero [et al.] // *Journal of Food Engineering*. – 2010. – 101. – P. 158-165. – ISSN 0260-8774. – Текст : электронный.

80. Guidi, L. How does chloroplast protect chlorophyll against excessive light / L. Guidi, M. Tattini, M. Landi // *Chlorophyll*. – 2017. – Vol. 21. – P. 21-36. – ISBN 978-953-51-3108-3. – Текст : электронный.

81. Henrich, V. Index Database: A database for remote sensing indices. [Electronic resource] / Henrich V., Brüser K. University of Bonn, Institute of Crop Science and Resource Conservation (INRES), 2017. – URL: <https://www.indexdatabase.de/db/i.php> (Дата обращения: 05.07.2022).

82. Hirri, A. Prediction of polyphenol fraction in virgin olive oil using mid-infrared attenuated total reflectance accessory-mid-infrared coupled with partial least squares regression / A. Hirri [et al.] // *International Journal of Food Properties*. – 2016. – 19(7). – P. 1504–1512. – ISSN 1094-2912. – Текст : электронный.

83. Hofman, P. J. Percentage of dry matter and oil content are not reliable indicators of fruit maturity or quality in late-harvested Hass' avocado / P. J. Hoffman, M. Jobin-Décor, J. Giles // *HortScience*. – Print ISSN: 0018-5345, Online ISSN: 2327-9834. – 2000. – Vol. 35. – №. 4. – P. 694-695. – Текст : электронный.

84. Hong, J. H. et al. Near-infrared diffuse reflectance spectroscopic analysis of the amounts of moisture, protein, starch, amylose, and tannin in buckwheat flours / J. H. Hong [et al.] // *Journal of nutritional science and vitaminology*. – 1996. – Vol. 42. – №. 4. – P. 359-366. – ISSN 0301-4800. – Текст : непосредственный.
85. Hopkins, D.W. Derivatives in spectroscopy / D. W. Hopkins // *Near-infrared Analysis*. – 2001. – Vol. 2(1). – P. 1-13. – ISBN 978-0-8493-7393-0. – Текст : непосредственный.
86. Hooijschuur, J. H. Raman spectroscopy for future planetary exploration: photodegradation, self-absorption and quantification of carotenoids in microorganisms and mineral matrices / J. H. Hooijschuur [et al.] // *Journal of Raman Spectroscopy*. – 2015. – Vol. 46. – №. 10. – P. 856-862. – Print ISSN: 0377-0486, Online ISSN: 1097-4555. – Текст : электронный.
87. How Does Chloroplast Protect Chlorophyll Against Excessive Light? / Ed. E. Jacob-Lopes. IntechOpen, London, UK. – 2017. – 130 p. – ISBN 978-953-51-3108-3. – Текст : электронный.
88. Huang, M. Prediction of color and moisture content for vegetable soybean during drying using hyperspectral imaging technology / M. Huang [et al.] // *Journal of Food Engineering*. – 2014. – 128. – P. 24-30. – ISSN 0260-8774. – Текст : электронный.
89. Hug-Iten, S. Staling of bread: Role of amylose and amylopectin and influence of starch-degrading enzymes / S. Hug-Iten, F. Escher, B. Conde-Petit // *Cereal Chemistry*. – 2003. – Vol. 80. – №. 6. – P. 654-661. – ISSN 0009-0352. – Текст : непосредственный.
90. Hurtado-Fernández, E. Avocado fruit—Persea Americana / E. Hurtado-Fernández, A. Fernández-Gutiérrez, A. Carrasco-Pancorbo // *Exotic fruits*. – Academic Press, 2018. – P. 37-48. – ISBN 978-0-12-803138-4. – Текст : электронный.
91. Jackman, P. Automatic segmentation of beef longissimus dorsi muscle and marbling by an adaptable algorithm / P. Jackman, D. W. Sun, P. Allen // *Meat Science*. – 2009. – Vol. 83. – №. 2. – P. 187-194. – ISSN 0309-1740. – Текст : электронный.

92. Jamshidi, B. Prediction of soluble solids in oranges using visible/near-infrared spectroscopy: effect of peel / B. Jamshidi [et al.] // *International Journal of Food Properties*. – 2014. – Vol. 17. – №. 7. – P. 1460-1468. – ISSN 1094-2912. – Текст : электронный.
93. Jin, H. Rapid and non-destructive determination of moisture content of peanut kernels using hyperspectral imaging technique / H. Jin, L. Li, J. Cheng // *Food Analytical Methods*. – 2015. – 8. – P. 2524-2532. – ISSN 1936-9751. – Текст : непосредственный.
94. Joe, A. A. F. Identification of spectral regions of the key components in the near infrared spectrum of wheat grain / A. A. F. Joe, A. Gopal // *2017 International Conference on Circuit, Power and Computing Technologies (ICCPCT)*. – IEEE, 2017. – P. 1-5. – ISBN: 978-1-5090-4968-4. – Текст : электронный.
95. Jung, S. Comparison of Kjeldahl and Dumas methods for determining protein contents of soybean products / S. Jung [et al.] // *Journal of the American Oil Chemists' Society*. – 2003. – Vol. 80. – №. 12. – P. 1169-1173. – ISSN 0003-021X. – Текст : непосредственный.
96. Katsumoto, Y. Modern pretreatment methods in NIR spectroscopy / Y. Katsumoto [et al.] // *Near-infrared Analysis*. – 2001. – №. 2. – P. 29-36. – ISBN 978-0-8493-7393-0. – Текст : непосредственный.
97. Kemps, B. Assessment of the quality parameters in grapes using vis/nir spectroscopy / B. Kemps [et al.] // *Biosystems Engineering*. – 2010. – 105. – P. 507–513. – Online ISSN: 1537-5129, Print ISSN: 1537-5110. – Текст : электронный.
98. Kemsley, E. K. Discrimination between *Coffea arabica* and *Coffea canephora* variant *robusta* beans using infrared spectroscopy / E. K. Kemsley, S. Ruault, R. H. Wilson // *Food Chemistry*. – 1995. – Т. 54. – P. 321-326. – ISSN 0308-8146. – Текст : электронный.
99. Kishor, K. Analysis of milk adulteration using mid-IR spectroscopy / K. Kishor, R. Thakur // *International Journal on Recent and Innovation Trends in Computing and Communication*. – 2015. – Vol. 3. – №. 10. – P. 5890-5895. – ISSN 2321-8169. – Текст : электронный.

100. Kuhnen, S. ATR-FTIR spectroscopy and chemometric analysis applied to discrimination of landrace maize flours produced in southern Brazil / S. Kuhnen [et al.] // *International Journal of Food Science & Technology*. – 2010. – Vol. 45. – №. 8. – P. 1673-1681. – ISSN 0950-5423. – Текст : электронный.

101. Kumagai, M. A near infrared spectroscopic discrimination of noodle flours using a principal-component analysis coupled with chemical information / M. Kumagai [et al.] // *Analytical sciences*. – 2002. – Vol. 18. – №. 10. – P. 1145-1150. – ISSN 0910-6340. – Текст : электронный.

102. Kunugi, M. Evolution of green plants accompanied changes in light-harvesting systems / M. Kunugi [et al.] // *Plant and Cell Physiology*. – 2016. – Vol. 57. – №. 6. – P. 1231-1243. – ISSN 0032-0781 (Print), ISSN 1471-9053 (Online). – Текст : электронный.

103. Kwon, Y. K. Rapid metabolic discrimination and prediction of dioscin content from African yam tubers using Fourier transform-infrared spectroscopy combined with multivariate analysis / Y. K. Kwon [et al.] // *Food Chemistry*. – 2015. – 166. – P. 389-396. – ISSN 0308-8146. – Текст : непосредственный.

104. Lee, J. S. Quality prediction of kiwifruit based on near infrared spectroscopy / J. S. Lee [et al.] // *Korean Journal of Horticultural Science and Technology*. – 2012. – 30. – P. 709–717. – ISSN 1226-8763(Print), ISSN 2465-8588(Online) – Текст : электронный.

105. Liu, Y. D. Appli3 peaches. In *Monitoring Food Safety, Agriculture, and Plant Health*; Bennedsen B.S., Chen Y.R., Meyer G.E., Senecal A.G., Tu S.I., Eds. / Y. D. Liu [et al.] // SPIE: Providence, USA. – 2004. – Vol. 5271. – P. 347–355. ISSN 1996-756X. – Текст : электронный.

106. Liu, Y. Nondestructive measurement of soluble solid content of navel orange fruit by visible-nir spectrometric technique with pls-r and pca-bpnn / Y. Liu, X. Sun, A. Ouyang // *LWT—Food Science and Technology*. – 2010. – 43. – P. 602–607. – ISSN 0023-6438. Текст : электронный.

107. Lohumi, S. A review of vibrational spectroscopic techniques for the detection of food authenticity and adulteration / S. Lohumi [et al.] // *Trends in Food Science & Technology*. – 2015. – Vol. 46. – №. 1. – P. 85-98. – ISSN 0924-2244. – Текст : электронный.
108. López, A. Crude protein content determination of potatoes by NIRS technology / A. López [et al.] // *Procedia Technology*. – 2013. – 8. – P. 488-492. – ISSN 2212-0173. – Текст : непосредственный.
109. Lu, J. Field detection of anthracnose crown rot in strawberry using spectroscopy technology / J. Lu [et al.] // *Computers and electronics in agriculture*. – 2017. – Vol. 135. – P. 289-299. – Print ISSN: 0168-1699, Online ISSN: 1872-7107. – Текст : электронный.
110. Lü, Q. Study of simplification of prediction model for kiwifruit firmness using near infrared spectroscopy / Q. Lü [et al.] // *Spectroscopy and Spectral Analysis*. – 2009. – 29. – P. 1768–1771. – ISSN 1000-0593. – Текст : электронный.
111. Luypaert, J. An evaluation of direct orthogonal signal correction and other pre-processing methods for the classification of clinical study lots of a dermatological cream / J. Luypaert [et al.] // *Journal of pharmaceutical and biomedical analysis*. – 2002. – Vol. 30. – №. 3. – P. 453-466. – ISSN 0731-7085. – Текст : непосредственный.
112. Lyman, D. J. FTIR-ATR analysis of brewed coffee: effect of roasting conditions / D. J. Lyman [et al.] // *Journal of Agricultural and Food Chemistry*. – 2003. – Vol. 51. – P. 3268-3272. – Online ISSN: 1520-5118, Print ISSN: 0021-8561. – Текст : электронный.
113. Lyu, Y. Identify the ripening stage of avocado by multispectral camera using semi-supervised learning on small dataset : дис. – 2019.
114. Maggio, R.M. Monitoring of fatty acid composition in virgin olive oil by Fourier transformed infrared spectroscopy coupled with partial least squares / R. M. Maggio [et al.] // *Food Chemistry*. – 2009. – 114(4). – P. 1549-1554. – ISSN 0308-8146. – Текст : непосредственный.

115. Magwaza, L. S. A review of destructive and non-destructive methods for determining avocado fruit maturity / L. S. Magwaza, S. Z. Tesfay // Food and Bioprocess Technology. – 2015. – Vol. 8. – №. 10. – P. 1995-2011. – ISSN 1935-5130. – Текст : электронный.

116. Mahesh, S. Comparison of partial least squares regression (PLSR) and principal components regression (PCR) methods for protein and hardness predictions using the near-infrared (NIR) hyperspectral images of bulk samples of Canadian wheat / S. Mahesh [et al.] // Food and Bioprocess Technology. – 2015. – 8. – P. 31-40. – ISSN 1935-5130. – Текст : электронный.

117. Mailer, R. J. Rapid evaluation of olive oil quality by NIR reflectance spectroscopy / R. J. Mailer // Journal of the American Oil Chemists' Society. – 2004. – 81(9). – P. 823-827. – ISSN 0003-021X. – Текст : непосредственный.

118. Marigheto, N. A. A comparison of mid-infrared and raman spectroscopies for the authentication of edible oils / N. A. Marigheto [et al.] // Journal of the American Oil Chemists' Society. – 1998. – 75(8). – P. 987-992. – ISSN 0003-021X. – Текст : непосредственный.

119. Marten, G. C. Near-infrared Reflectance Spectroscopy (NIRS): Analysis of Forage Quality / G.C. Marten, J. S. Shenk, J. E. Barton // Washington, DC: United States Department of Agriculture. – 1989. – 110 p. – ISSN 0065-4612. – Текст : непосредственный.

120. Martens, H. Multivariate calibration. I. Concepts and distinctions / H. Martens, T. Næs // TrAC Trends in Analytical Chemistry. – 1984. – Vol. 3. – №. 8. – P. 204-210. – ISSN 0165-9936. – Текст : непосредственный.

121. Martens, H. Multivariate Calibration / H. Martens, T. Næs // 1 edition. – Chichester: John Wiley & Sons. – 1992. – 419 p. – ISBN 978-0471930471. – Текст : электронный.

122. Massart, D. L. Chemometrics: A Textbook / D. L. Massart [et al.] // Vol. 2. Amsterdam: Elsevier. – 1988. – 299 p. – ISBN 978-0-0808-6829-5. – Текст : непосредственный.

123. Matsuda, O. Hyperspectral imaging techniques for rapid identification of Arabidopsis mutants with altered leaf pigment status / O. Matsuda [et al.] // *Plant and Cell Physiology*. – 2012. – Т. 53. – №. 6. – С. 1154-1170. – ISSN (Print): 0032-0781, ISSN (Online) 1471-9053. – Текст : электронный.

124. McGlone, V. A. Firmness, dry-matter and soluble-solids assessment of postharvest kiwifruit by nir spectroscopy / V. A. McGlone, S. Kawano // *Postharvest Biology and Technology*. – 1998. – 13. – P. 131–141. – Print ISSN: 0925-5214, Online ISSN: 1873-2356. – Текст : электронный.

125. Mendes, E., Duarte, N. Mid-infrared spectroscopy as a valuable tool to tackle food analysis: A literature review on coffee, dairies, honey, olive oil and wine / E. Mendes, N. Duarte // *Foods*. – 2021. – Vol. 10. – №. 2. – P. 477. – ISSN 2304-8158. – Текст : электронный.

126. Mendoza, F. A. Evaluation of Canning quality traits in black beans (*Phaseolus vulgaris* L.) by visible/near-infrared spectroscopy / F. A. Mendoza [et al.] // *Food and Bioprocess Technology*. – 2014. – 7. – P. 2666-2678. – ISSN 1935-5130. – Текст : электронный.

127. Merzlyak, M. N. Reflectance spectral features and non-destructive estimation of chlorophyll, carotenoid and anthocyanin content in apple fruit / M. N. Merzlyak, A. E. Solovchenko, A. A. Gitelson // *Postharvest biology and technology*. – 2003. – Vol. 27. – №. 2. – P. 197-211. – Print ISSN: 0925-5214, Online ISSN: 1873-2356. – Текст : электронный.

128. Metlenkin, D. A. Identification of coffee beans using ftir-spectroscopy and multivariate analysis / D.A. Metlenkin [et al.] // *News of institutes of higher education. Food technology*. – 2021. – №5–6. – P.92-97. – ISSN 0579-3009. – Текст : непосредственный.

129. Metlenkin, D.A. Non-destructive identification of defects and classification of Hass avocado fruits with the use of a hyperspectral image / D. A. Metlenkin [et al.] // *Agronomy Research*. – 2022. – Vol. 20. – No. 2. – P. 326-340. – ISSN 1406-894X. – Текст : электронный.

130. Milenković, S. M. The identification of chlorophyll and its derivatives in the pigment mixtures: hplc-chromatography, visible and mass spectroscopy studies / S. M. Milenković [et al.] // *Advanced technologies*. – 2012. – Vol. 1. – №. 1. – P. 16-24. – ISSN (Print): 2406-2979, ISSN (Online) 2406-3037. – Текст : электронный.

131. Moghimi, A. Vis/nir spectroscopy and chemometrics for the prediction of soluble solids content and acidity (ph) of kiwifruit / A. Moghimi [et al.] // *Biosystems Engineering*. – 2010. – 106. – P. 295–302. – Online ISSN: 1537-5129, Print ISSN: 1537-5110. – Текст : электронный.

132. Monteiro, S. T. Prediction of sweetness and amino acid content in soybean crops from hyperspectral imagery / S. T. Monteiro [et al.] // *ISPRS Journal of Photogrammetry and Remote Sensing*. – 2007. – 62. – P. 2-12. – ISSN 0924-2716. – Текст : электронный.

133. Næs, T. A user-friendly guide to multivariate calibration and classification / T. Næs [et al.] // Chichester, UK: NIR Publications. – 2002. – 344 p. – ISBN 978-1-906715-25-0 – Текст : непосредственный.

134. Næs, T. Fitting, prediction testing, cross-validation or leverage correction? / T. Næs, T. Isaksson // *NIR News*. – 1991. – Vol. 2. – P. 10- 13. – ISSN 0960-3360. – Текст : непосредственный.

135. Nagata, M. Bruise detection using nir hyperspectral imaging for strawberry (*fragaria x ananassa duch.*) / M. Nagata, J. G. Tallada, T. Kobayashi // *Environmental Control in Biology*. – 2006. – 44. – P.133–142. – ISSN 1880554X. – Текст : электронный.

136. Nenadis, N. Perspective of vibrational spectroscopy analytical methods in on-field/official control of olives and virgin olive oil / N. Nenadis, M. Z. Tsimidou // *European Journal of Lipid Science and Technology*. – 2017. – 119(1). – article 1600148. – ISSN: 1438-9312 (online). – Текст : электронный.

137. Nie, P. Fast determination of boiling time of yardlong bean using visible and near infrared spectroscopy and chemometrics / P. Nie [et al.] // *Journal of Food Engineering*. – 2012. – Vol. 109. – №. 1. – P. 155-161. – ISSN 0260-8774. – Текст : непосредственный.

138. Niu, X. Y. Nondestructive discrimination of strawberry varieties by nir and bp-ann / X. Y. Niu [et al.] // *Spectroscopy and Spectral Analysis*. – 2012. – 32. – P. 2095-2099. – ISSN 1000-0593. – Текст : электронный.

139. Nurrulhidayah, A. F. Analysis of chicken fat as adulterant in butter using Fourier transform infrared spectroscopy and chemometrics / A. F. Nurrulhidayah [et al.] // *Grasas y Aceites*. – 2013. – Vol. 64. – №. 4. – С. 349-355. – ISSN 0017-3495. – Текст : электронный.

140. Nurrulhidayah, A. F. Application of FTIR-ATR spectroscopy coupled with multivariate analysis for rapid estimation of butter adulteration / A. F. Nurrulhidayah [et al.] // *Journal of Oleo Science*. – 2013. – Vol. 62. – №. 8. – P. 555-562. – ISSN 1345-8957. – Текст : электронный.

141. Nurrulhidayah, A. F. Authentication analysis of butter from beef fat using Fourier Transform Infrared (FTIR) spectroscopy coupled with chemometrics / A. F. Nurrulhidayah [et al.] // *International Food Research Journal*. – 2013. – Vol. 20. – №. 3. – P. 1383-1388. – ISSN 2231-7546. – Текст : электронный.

142. Obenland, D. Influence of maturity and ripening on aroma volatiles and flavor in ‘Hass’ avocado / D. Obenland [et al.] // *Postharvest Biology and Technology*. – 2012. – Vol. 71. – P. 41-50. – Print ISSN: 0925-5214, Online ISSN: 1873-2356. – Текст : электронный.

143. Ochoa-Ascencio, S. Modelling the transient effect of 1-MCP on ‘Hass’ avocado softening: A Mexican comparative study / S. Ochoa-Ascencio, M. L. Hertog, B. M. Nicolai // *Postharvest Biology and Technology*. – 2009. – Vol. 51. – №. 1. – P. 62-72. – Print ISSN: 0925-5214, Online ISSN: 1873-2356. – Текст : электронный.

144. Oliveri, P. Multivariate class modeling for the verification of food-authenticity claims / P. Oliveri, G. Downey // *TrAC Trends in Analytical Chemistry*. – 2012. – Vol. 35. – P. 74-86. – ISSN 0165-9936. – Текст : электронный.

145. Ollinger, S. V. Sources of variability in canopy reflectance and the convergent properties of plants / S. V. Ollinger // *New Phytologist*. – 2011. – Vol. 189. – №. 2. – P. 375-394. – Print ISSN: 0028-646X, Online ISSN: 1469-8137. – Текст : электронный.

146. Omar, A. F. Spectroscopic profiling of soluble solids content and acidity of intact grape, lime, and star fruit / A. F. Omar // *Sensor Review*. – 2013. – 33. – P. 238–245. – Print ISSN: 0260-2288. – Текст : электронный.

147. Osborne, B.G. Near-infrared spectroscopy in food analysis / B. G. Osborne // *Encyclopedia of analytical chemistry: applications, theory and instrumentation*. – 2006. – ISBN 978-0-4719-7670-7 (в обл.). – ISBN 978-0-4700-2731-8 (e-book). – Текст : непосредственный.

148. Osborne, B. G. Spectroscopy with Application in Food and Beverage Analysis / B.G. Osborne [et al.] // Singapore: Longman Scientific and Technical. – 1993. – 227 p. – ISBN 978-0-4702-2128-0. – Текст : непосредственный.

149. Parodi, G. Correlation of oil content, dry matter and pulp moisture as harvest indicators in Hass avocado fruit (*Persea americana* Mill) grown under two conditions of orchards in Chincha-Peru / G. Parodi, M. Sanchez, W. Daga // *Proceedings VI World Avocado Congress, Chile*. – 2007. – P. 12-16. – ISBN 978-956-17-0413-8. – Текст : электронный.

150. Pathare, P. B. Colour measurement and analysis in fresh and processed foods: a review / P. B. Pathare, U. L. Opara, F. A. J. Al-Said // *Food and Bioprocess Technology*. – 2013. – 6. – P. 36-60. – ISSN 1935-5130. – Текст : электронный.

151. Pomerantsev, A. L. Concept and role of extreme objects in PCA/SIMCA / A. L. Pomerantsev, O. Y. Rodionova // *Journal of Chemometrics*. – 2014. – Vol. 28. – №. 5. – P. 429-438. – ISSN 0886-9383. – Текст : электронный.

152. Pomerantsev, A. L. Multiclass partial least squares discriminant analysis: Taking the right way—A critical tutorial / A. L. Pomerantsev, O. Y. Rodionova // *Journal of Chemometrics*. – 2018. – Vol. 32. – №. 8. – e3030. – ISSN 0886-9383. – Текст : электронный.

153. Posom, J. Predicting Marian Plum Fruit Quality without Environmental Condition Impact by Handheld Visible–Near-Infrared Spectroscopy / J. Posom [et al.] // *ACS omega*. – 2020. – Vol. 5. – №. 43. – P. 27909-27921. – ISSN 2470-1343. – Текст : электронный.

154. Pu, Y. Y. Recent progress of hyperspectral imaging on quality and safety inspection of fruits and vegetables: a review / Y. Y. Pu, Y. Z. Feng, D. W. Sun // *Comprehensive Reviews in Food Science and Food Safety*. – 2015. – Vol. 14. – №. 2. – P. 176-188. – ISSN 1541-4337. – Текст : электронный.

155. Qin, P. Nutritional composition and flavonoids content of flour from different buckwheat cultivars / P. Qin [et al.] // *International Journal of Food Science & Technology*. – 2010. – Vol. 45. – №. 5. – P. 951-958. – ISSN 0950-5423. – Текст : электронный.

156. Rady, A. M. The potential use of visible/near infrared spectroscopy and hyperspectral imaging to predict processing-related constituents of potatoes / A. M. Rady [et al.] // *Journal of Food Engineering*. – 2014. – 135. – P. 11-25. – ISSN 0260-8774. – Текст : непосредственный.

157. Radzevičius, A. Determination of tomato quality attributes using near infrared spectroscopy and reference analysis / A. Radzevičius [et al.] // *Zemdirbyste-Agriculture*. – 2016. – 103(1). – P. 91-98. – ISSN 1392-3196. – Текст : электронный.

158. Ricci, A. Application of Fourier Transform Infrared (FTIR) Spectroscopy in the Characterization of Tannins / A. Ricci [et al.] // *Applied Spectroscopy Reviews*. – 2015. – Vol. 50 (5). – P. 407-442. – Online ISSN: 1520-569X, Print ISSN: 0570-4928. – Текст : электронный.

159. Riyanto, S. Analysis Of Palm Oil As Oil Adulterant In Olive And Pumpkin Seed Oils In Ternary Mixture Systems Using Ftir Spectroscopy And Chemometrics / S. Riyanto, S. Martono, A. Rohman // *International Journal of Applied Pharmaceutics*. – 2019. – 11(5). – P. 210-215. – ISSN 0975-7058. – Текст : электронный.

160. Sacks, E. J. Genetic and environmental variation for tomato flesh color in a population of modern breeding lines / E. J. Sacks, D. M. Francis // *Journal of the American*

Society for Horticultural Science. – 2001. – Vol. 126. – №. 2. – P. 221-226. – ISSN 0003-1062. – Текст : электронный.

161. Safar, M. Characterization of edible oils, butters and margarines by Fourier transform infrared spectroscopy with attenuated total reflectance / M. Safar [et al.] // Journal of the American Oil Chemists' Society. – 1994. – Vol. 71. – P. 371-377. – ISSN 0003-021X. – Текст : непосредственный.

162. Samyn, P. Quality and statistical classification of Brazilian vegetable oils using mid-infrared and Raman spectroscopy / P. Samyn [et al.] // Applied Spectroscopy. – 2012. – 66(5) – P. 552-565. – ISSN 0003-7028. – Текст : электронный.

163. Sanchez, M. T. Non-destructive characterization and quality control of intact strawberries based on nir spectral data / M. T. Sanchez [et al.] // Journal of Food Engineering. – 2012. – 110. – P. 102–108. – ISSN 0260-8774. – Текст : электронный.

164. Shao, Y. Nondestructive measurement of the internal quality of bayberry juice using vis/nir spectroscopy / Y. Shao, Y. He // Journal of Food Engineering. – 2007. – 79. – P. 1015-1019. – ISSN 0260-8774. – Текст : электронный.

165. Sun, D. W. (ed.). Hyperspectral imaging for food quality analysis and control. – Elsevier, 2010. – ISBN 978-0-1281-0223-7 (в обл.). – ISBN 978-0-0808-8628-2 (e-book). – Текст : электронный.

166. Torres, I. An overview of regression methods in hyperspectral and multispectral imaging / I. Torres, J. M. Amigo // Data Handling in Science and Technology. – 2019. – Vol. 32. – P. 205-230. – ISSN (Print): 0922-3487, ISSN (Online) 2212-1250. – Текст : электронный.

167. Uarrota, V. G. Metabolomics combined with chemometric tools (PCA, HCA, PLS-DA and SVM) for screening cassava (*Manihot esculenta* Crantz) roots during postharvest physiological deterioration / V. G. Uarrota [et al.] // Food Chemistry. – 2014. – 161. – P. 67-78. – ISSN 0308-8146. – Текст : непосредственный.

168. Uncu, O. Prediction of various chemical parameters of olive oils with Fourier transform infrared spectroscopy / O. Uncu, B. Ozen // *LWT—Food Science and Technology*. – 2015. – 63(2). – P. 978-984. – ISSN 0023-6438. – Текст : электронный.

169. Voitsekhovskaja, O. V. Chlorophyll b in angiosperms: functions in photosynthesis, signaling and ontogenetic regulation / O.V. Voitsekhovskaja, E. V. Tyutereva // *Journal of Plant Physiology*. – 2015. – Vol. 189. – P. 51-64. – ISSN 0176-1617. – Текст : электронный.

170. Wang, F. Research of Determination Method of Starch and Protein Content in Buckwheat by Mid-Infrared Spectroscopy / F. Wang [et al.] // *International Conference on Computer and Computing Technologies in Agriculture*. – Springer, Berlin, Heidelberg, 2012. – P. 248-254. – ISSN 1861-2288. – Текст : непосредственный.

171. Wang, H. Fruit Quality Evaluation Using Spectroscopy Technology: A Review / H. Wang [et al.] // *Sensors*. – 2015. – 15(5). – P. 11889-11927. – Online ISSN: 1424-8220. – Текст : электронный.

172. Wang, L. Use of hyperspectral imaging to discriminate the variety and quality of rice / L. Wang [et al.] // *Food Analytical Methods*. – 2015. – 8. – P. 515-523. – ISSN 1936-9751. – Текст : непосредственный.

173. Wedding, B. B. Non-destructive prediction of ‘Hass’ avocado dry matter via FT-NIR spectroscopy / B. B. Wedding [et al.] // *Journal of the Science of Food and Agriculture*. – 2011. – Vol. 91. – №. 2. – P. 233-238. – ISSN 2575-7954. – Текст : электронный.

174. Wei, F. Chemical changes in the components of coffee beans during roasting / F. Wei, M. Tanokura // *Coffee in health and disease prevention*. – Academic Press, 2015. – P. 83-91. – ISBN 978-0-12-409517-5. – Текст : электронный.

175. Wetzel, D.L. Charalambous G. Instrumental Methods in Food and Beverage Analysis / D. L. Wetzel // *Amsterdam: Elsevier*. – 1998. – Vol. 39(1). – 629 p. – ISBN 978-0-0805-3475-6 – Текст : непосредственный.

176. Wetzel D. L. Near-infrared reflectance analysis / D. L. Wetzel // *Analytical chemistry*. – 1983. – Vol. 55. – №. 12. – P. 1165A-1176A. – ISBN 978-0-4708-8757-8. – Текст : непосредственный.

177. Williams, P. Near-infrared Technology in the Agricultural and Food Industries / P. Williams, K. Norris // 2nd edition. – St Paul, MN: American Association of Cereal Chemists. –2001. – 312 p. – ISBN 978-1-891127-24-3. – Текст : непосредственный.

178. Williams, P. C. Comparison of commercial near-infrared transmittance and reflectance instruments for analysis of whole grains and seeds / P. C. Williams, D. C. Sobering // *Journal of Near-infrared Spectroscopy*. – 1993. – Vol. 1. – P. 25-32. – ISSN 0967-0335. – Текст : непосредственный.

179. Wold, S. Chemometrics; what do we mean with it, and what do we want from it? / S. Wold // *Chemometrics and intelligent laboratory systems*. – 1995. – Vol. 30. – №. 1. – P. 109-115. – ISSN 0169-7439. – Текст : непосредственный.

180. Wold, S. Orthogonal signal correction of near-infrared spectra / S. Wold [et al.] // *Chemometrics Intelligent Laboratory Systems*. – 1998. – №. 44. – P. 175- 185. – ISSN 0169-7439. – Текст : непосредственный.

181. Wu, D. Fast discrimination of juicy peach varieties by Vis/NIR spectroscopy based on bayesian-sda and pca / D. Wu, Y. He, Y. D. Bao // *Lecture Notes in Computer Science*. – 2006. – 4113. – P. 931-936. – Electronic ISSN 1611-3349, Print ISSN 0302-9743. – Текст : электронный.

182. Xia, J. Approach to nondestructive measurement of vitamin c content of orange with near-infrared spectroscopy treated by wavelet transform / J. Xia [et al.] // *Transactions of the Chinese Society of Agricultural Engineering*. – 2007. – 23. – P. 170–174. – ISSN 1002-6819. – Текст : электронный.

183. Xing, J. Comparison of short-wavelength infrared (SWIR) hyperspectral imaging system with an FT-NIR spectrophotometer for predicting alpha-amylase activities in individual Canadian Western Red Spring (CWRS) wheat kernels / J. Xing [et al.] // *Biosystems*

Engineering. – 2011. – 108(4). – P. 303-310. – Online ISSN: 1537-5129, Print ISSN: 1537-5110. Текст : непосредственный.

184. Yan, R. Discrimination of strawberries varieties based on characteristic spectrum / R. Yan [et al.] // Transactions of the Chinese Society for Agricultural Machinery. – 2013. – 44. – P. 182-186. – ISSN 1000-1298. – Текст : электронный.

185. Yilmaz, H. Ö. Buckwheat: A useful food and its effects on human health / H. Ö. Yilmaz, N. Y. Ayhan, Ç. S. Meriç // Current Nutrition & Food Science. – 2020. – Vol. 16. – №. 1. – P. 29-34. – ISSN (Print): 1573-4013, ISSN (Online): 2212-3881. – Текст : электронный.

186. Yu, J. J. Nondestructive test on predicting sugar content and valid acidity of mango by spectroscopy technology / J. J. Yu, Y. He, Y. D. Bao // Spectroscopy and Spectral Analysis. – 2008. – 28. – P. 2839–2842. – ISSN 1000-0593. – Текст : электронный.

187. Zhang, S. Comparison of modeling methods of fresh jujube soluble solids measurement by nir spectroscopy / S. Zhang [et al.] // Transactions of the Chinese Society for Agricultural Machinery. – 2012. – 43. – P. 108–112. – ISSN 1000-1298. – Текст : электронный.

188. Zontov, Y. V. PLS-DA—A MATLAB GUI tool for hard and soft approaches to partial least squares discriminant analysis / Y. V. Zontov [et al.] // Chemometrics and Intelligent Laboratory Systems. – 2020. – Vol. 203. – P. 104064. – ISSN 0169-7439. – Текст : электронный.

Приложение А (справочное)

Свидетельство о государственной регистрации базы данных

РОССИЙСКАЯ ФЕДЕРАЦИЯ

RU**2023620795**

ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ
(12) ГОСУДАРСТВЕННАЯ РЕГИСТРАЦИЯ БАЗЫ ДАННЫХ

Номер регистрации (свидетельства):

[2023620795](#)

Дата регистрации: **06.03.2023**

Номер и дата поступления заявки:

2023620470 20.02.2023

Дата публикации: [06.03.2023](#)

Контактные реквизиты:

нет

Авторы:

**Платов Юрий Тихонович (RU),
Метленкин Дмитрий Андреевич (RU),
Платова Ранса Абдулгафаровна (RU)**

Правообладатель:

**Федеральное государственное бюджетное
образовательное учреждение высшего образования
"Российский экономический университет имени
Г.В. Плеханова" (RU)**

Название базы данных:

Спектральные характеристики по данным гиперспектральных изображений плодов авокадо, различающиеся по влажности и наличию дефектов

Реферат:

База данных включает спектральные характеристики плодов авокадо, различающиеся по влажности и наличию дефектов. Спектральные характеристики плодов авокадо получены по данным гиперспектральных изображений в видимом и ближнем инфракрасном диапазоне с использованием камеры Specim IQ с последующей сегментацией изображений и отбором спектральных сигнатур. В базе данных приведены значения показателей плодов авокадо: страна происхождения, год урожая, сорт, влажность (MS,%), сухой остаток (DM,%), дефекты (незначительные и критические), коэффициенты отражения в диапазоне длин волн от 400 до 1000 нм (с шагом 3 нм). Область применения – разработка градуировочных моделей прогнозирования влажности и сухого остатка и классификационной модели градации по наличию дефектов плодов авокадо. Тип ЭВМ: IBM PC-совмест. ПК; ОС: Windows 10 и выше.

Вид и версия системы управления базой данных: Excel 2010

Объем базы данных: 1,5 МБ

Приложение Б (обязательное)**Справка о внедрении результатов исследования в учебный процесс**

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФГБОУ ВО «РЭУ им. Г.В. Плеханова»

СПРАВКА

о внедрении результатов диссертационных исследований

Подтверждаем, что результаты диссертационной работы Метленкина Дмитрия Андреевича «Разработка методических подходов применения оптической спектроскопии и гиперспектрального изображения для идентификации и контроля качества пищевых продуктов» применяются в образовательном процессе ФГБОУ ВО «РЭУ им. Г.В. Плеханова» с 2019 года.

Полученные в процессе диссертационного исследования результаты использовались при преподавании дисциплин: «Таможенная экспертиза зерно-мучных и плодоовощных товаров», «Идентификация и обнаружение фальсификации товаров» у бакалавров и «Методология научных исследований в товароведении» у магистров на кафедре товароведения и товарной экспертизы.

Разработанные автором методические подходы применения оптической спектроскопии и гиперспектрального изображения для идентификации и контроля качества пищевых продуктов используются при подготовке курсовых и выпускных квалификационных работ бакалавров и магистров в рамках учебного процесса по направлениям подготовки «Товарный менеджмент и экспертиза», «Пищевые инновации и продакт-дизайн» ФГБОУ ВО «РЭУ им. Г.В. Плеханова»

Заведующий кафедрой
товароведения и товарной экспертизы



П.И. Токарев

Приложение В (обязательное)

Справка о внедрении результатов исследования в лаборатории качества Qlab сети «Азбуки Вкуса»



от 14.03.2023г.

СПРАВКА

о внедрении результатов диссертационного исследования

Подтверждаем, что результаты диссертационного исследования Метленкина Дмитрия Андреевича «Разработка методических подходов применения оптической спектроскопии и гиперспектрального изображения для идентификации и контроля качества пищевых продуктов» имеют практическую значимость и применялись в лаборатории качества Qlab сети «Азбуки Вкуса», в том числе:

1) Автором диссертационного исследования и сотрудниками лаборатории качества «Азбуки Вкуса» совместно были разработаны и апробированы компьютерные модели для определения влажности и сухого остатка плодов авокадо сорта Хасс;

2) По результатам совместной деятельности подготовлена научная статья «Сортировка плодов авокадо по гиперспектральным изображениям», которая находится на рецензировании в журнале «Пищевые системы» (RSCI).

Начальник лаборатории Qlab



О.В. Федосеенко