



Продовольственная и  
сельскохозяйственная организация  
Объединенных Наций

1

ОБУЧАЮЩИЕ  
МАТЕРИАЛЫ

# СПЕКТРОСКОПИЯ ПОЧВ ОБУЧАЮЩИЕ МАТЕРИАЛЫ

1

Руководство по анализу почв с использованием  
видимой и ближней инфракрасной (vis-NIR) и  
средней инфракрасной (MIR) спектроскопии





# СПЕКТРОСКОПИЯ ПОЧВ ОБУЧАЮЩИЕ МАТЕРИАЛЫ

**1** | Руководство по анализу почв с использованием видимой и ближней инфракрасной (vis-NIR) и средней инфракрасной (MIR) спектроскопии

**Юфэн Гэ**

Университет Небраски

**Александр Ваду**

Университет Сиднея

**Ий Пэн**

Глобальное почвенное партнерство, ФАО

Обязательная ссылка

ФАО. 2022. *Руководство по анализу почв с использованием видимой и ближней инфракрасной (vis-NIR) и средней инфракрасной (MIR) спектроскопии*. Рим, ФАО  
<https://doi.org/10.4060/cb9005ru>

Используемые обозначения и представление материала в настоящем информационном продукте не означают выражения какого-либо мнения со стороны Продовольственной и сельскохозяйственной организации Объединенных Наций (ФАО) относительно правового статуса или уровня развития той или иной страны, территории, города или района, или их властей, или относительно делимитации их границ или рубежей. Упоминание конкретных компаний или продуктов определенных производителей, независимо от того, запатентованы они или нет, не означает, что ФАО одобряет или рекомендует их, отдавая им предпочтение перед другими компаниями или продуктами аналогичного характера, которые в тексте не упоминаются.

Мнения, выраженные в настоящем информационном продукте, являются мнениями автора (авторов) и не обязательно отражают точку зрения или политику ФАО.

ISBN 978-92-5-136892-3

© ФАО, 2022



Некоторые права защищены. Настоящая работа предоставляется в соответствии с лицензией Creative Commons “С указанием авторства – Некоммерческая - С сохранением условий 3.0 НПО” (CC BY-NC-SA 3.0 IGO; <https://creativecommons.org/licenses/by-nc-sa/3.0/igo/deed.ru>).

Согласно условиям данной лицензии настоящую работу можно копировать, распространять и адаптировать в некоммерческих целях при условии надлежащего указания авторства. При любом использовании данной работы не должно быть никаких указаний на то, что ФАО поддерживает какую-либо организацию, продукты или услуги. Использование логотипа ФАО не разрешено. В случае адаптации работы она должна быть лицензирована на условиях аналогичной или равнозначной лицензии Creative Commons. В случае перевода данной работы, вместе с обязательной ссылкой на источник, в него должна быть включена следующая оговорка: «Данный перевод не был выполнен Продовольственной и сельскохозяйственной организацией Объединенных Наций (ФАО). ФАО не несет ответственности за содержание или точность данного перевода. Достоверной редакцией является издание на [указать язык оригинала] языке».

Возникающие в связи с настоящей лицензией споры, которые не могут урегулированы по обоюдному согласию, должны разрешаться через посредничество и арбитражное разбирательство в соответствии с положениями Статьи 8 лицензии, если в ней не оговорено иное. Посредничество осуществляется в соответствии с “Правилами о посредничестве” Всемирной организации интеллектуальной собственности <http://www.wipo.int/amc/ru/mediation/rules/index.html>, а любое арбитражное разбирательство должно производиться в соответствии с “Арбитражным регламентом” Комиссии Организации Объединенных

Наций по праву международной торговли (ЮНСИТРАЛ).

Материалы третьих лиц. Пользователи, желающие повторно использовать материал из данной работы, авторство которого принадлежит третьей стороне, например, таблицы, рисунки или изображения, отвечают за то, чтобы установить, требуется ли разрешение на такое повторное использование, а также за получение разрешения от правообладателя. Удовлетворение исков, поданных в результате нарушения прав в отношении той или иной составляющей части, авторские права на которую принадлежат третьей стороне, лежит исключительно на пользователе.

Продажа, права и лицензирование. Информационные продукты ФАО размещаются на веб-сайте ФАО ([www.fao.org/publications](http://www.fao.org/publications)); желающие приобрести информационные продукты ФАО могут обращаться по адресу: [publications-sales@fao.org](mailto:publications-sales@fao.org). По вопросам коммерческого использования следует обращаться по адресу: [www.fao.org/contact-us/licence-request](http://www.fao.org/contact-us/licence-request). За справками по вопросам прав и лицензирования следует обращаться по адресу: [copyright@fao.org](mailto:copyright@fao.org).

Фото на обложке: © Университет Небраски

# Оглавление

<b>Составители</b>	<b>V</b>
<b>Благодарности</b>	<b>V</b>
<b>Резюме</b>	<b>V</b>
<b>1   Введение</b>	<b>1</b>
<b>2   Основы использования vis-NIR и MIR излучения для анализа почв</b>	<b>1</b>
2.1 Спектр электромагнитного излучения, длина волны и волновое число	1
2.2 Спектры отражения и поглощения	1
2.3 Основные полосы поглощения в MIR, обертоны и комбинации в NIR	2
2.4 Свойства почвы, которые могут быть прямо или косвенно оценены с помощью спектров в vis-NIR и MIR диапазонах	2
2.5 Преимущества анализа почв с применением vis-NIR и MIR спектроскопии	3
<b>3   Процедуры анализа почв с применением vis-NIR и MIR спектроскопии</b>	<b>4</b>
3.1 Подготовка образцов	5
3.2 Сканирование спектров	5
3.3 Предварительная обработка спектров	6
3.4 Обучение и тестирование модели	7
3.5 Методы обучения модели	9
3.6 Оценка модели	10
<b>4   Спектральные библиотеки почв в vis-NIR и MIR диапазонах в региональном, континентальном и глобальном масштабах: актуальность, преимущества и ограничения</b>	<b>14</b>
<b>5   Приборы для сканирования почвы в vis-NIR и MIR диапазонах</b>	<b>17</b>
<b>6   Заключительные замечания</b>	<b>17</b>
<b>Библиография</b>	<b>18</b>

# Рисунки

1. Спектры почвы в видимом и БИК-диапазонах (vis-NIR): спектр отражения (слева) и поглощения (справа)	2
2. Рабочий процесс vis-NIR и MIR спектроскопии для анализа почв	4
3. Примеры спектров отражения Vis-NIR из экспериментального набора данных (n=201) с различными методами предварительной обработки спектра: (A) исходные спектры отражения; (B) спектры поглощения; (C) спектры после стандартного нормального изменения SNV; (D) спектры после мультипликативной коррекции сигнала MSC; (E) спектры после удаления континуума; и (F) спектры первой производной после фильтрации Савицкого-Голея. Синие линии - средние спектры, рассчитанные по всем образцам	7
4. Иллюстрация переобучения модели с частичной регрессией наименьших квадратов	8
5. Иллюстрация МГК (PCA) на спектральных данных почвенной MIR-спектроскопии: (A) данные почв в MIR диапазоне в спектральном пространстве; (B) график осыпи первых 10 основных компонент и процентная дисперсия, приходящаяся на каждую компоненту; и (C) попарные диаграммы рассеяния первых трех компонент	10
6. Смоделированные данные, показывающие сравнение предсказанных значений в vis-NIR или MIR диапазоне свойств почвы с измеренными в лаборатории свойствами на диаграммах рассеяния. Приведены различные уровни производительности модели и приведены показатели их оценки	12
7. Диаграммы рассеяния предсказанных значений в vis-NIR диапазоне и измеренных в лаборатории свойств почвы для тестового набора (n = 60) с использованием частичной регрессии наименьших квадратов. Пунктирная серая линия - линия равенства между наблюдаемым и прогнозируемым, синяя линия – линия тренда	12
8. Диаграммы рассеяния предсказанных MIR и измеренных в лаборатории свойств почвы для тестового набора (n = 108) с использованием частичной регрессии наименьших квадратов	13

# Таблицы

1. Результаты тестирования трех почвенных свойств с моделированием PLSR и SVR для значений в vis-NIR и MIR диапазонах	13
2. Краткое описание опубликованных почвенных спектральных библиотек в vis-NIR и MIR диапазонах в национальном, континентальном и глобальном масштабах.	15

# Составители

## Общая координация

Рональд Варгас Рохас (ФАО-ГПП)

## Авторы

Юфэн Гэ (Университет Небраски, США)

Александр Ваду (Сиднейский

университет, Австралия)

Йи Пэн (ФАО-ГПП)

## Редактирование и публикация

Филиппо Бенедетти (ФАО-ГПП)

Маттео Сала (ФАО-ГПП)

Тасним Алсиддиг (ФАО-ГПП)

## Перевод

Ольга Якименко (МГУ имени М.В. Ломоносова)

Мария Конюшкова (ФАО-ГПП)

## Рецензенты

Будиман Минасны (Сиднейский университет,

Австралия)

Хосе Александр Мело Дематте (Университет Сан-

Паулу, Бразилия)

# Благодарности

“Руководство по анализу почвы с использованием видимой и ближней инфракрасной (vis-NIR) и средней инфракрасной (MIR) спектроскопии” является первым учебным материалом по теме спектроскопии почв для начинающих, разработанным по инициативе Глобальной сети почвенных лабораторий по спектроскопии почв (GLOSOLAN-Spec) Глобального почвенного партнерства ФАО. Это результат сотрудничества экспертов в почвенной спектроскопии из разных институтов по всему миру. Этот документ представляет собой введение в использование почвенной спектроскопии для анализа почв, что позволяет читателям понять основы и базовые процедуры использования этой технологии для анализа почв. Команда GLOSOLAN-Spec и авторы пособия особенно благодарны проекту Всемирного банка “Использование технологий для модернизации сельского хозяйства Узбекистана”, финансируемому Корейским трастовым фондом Зеленого роста, за поддержку в подготовке этого документа, имеющего целью распространение современных технологий спектроскопического анализа почв во всех регионах мира. В заключение авторы хотели бы поблагодарить Лабораторию почвенных исследований Келлога Службы охраны природных ресурсов Министерства сельского хозяйства США (USDA-NRCS) за предоставление наборов данных почвенных спектров в vis-NIR и MIR диапазонах для этого пособия.

# Резюме

Спектроскопия отражения в видимом и ближнем инфракрасном (vis-NIR) и среднем инфракрасном (MIR) диапазонах появилась и развивалась как важный метод количественного анализа почв, который потенциально может стать альтернативой традиционному лабораторному анализу некоторых почвенных свойств методами «мокрой» химии. Vis-NIR и MIR предпочтительны из-за их скорости, низкой стоимости и неразрушаемости проб при анализе, но они требуют нового набора навыков у персонала лаборатории и практиков. Это вводное руководство предназначено для начинающих специалистов в этой области, которые хотят использовать vis-NIR и MIR спектроскопию для анализа почв. Учебное пособие охватывает следующие темы: (1) основы vis-NIR и MIR излучения и их взаимодействия с почвой (2) общие лабораторные процедуры для анализа почвы в vis-NIR и MIR диапазонах с акцентом на получение спектра, предварительную обработку спектра, обучение и тестирование модели, частичную регрессию наименьших квадратов и оценку эффективности модели и (3) библиотеки почвенных vis-NIR и MIR спектров в региональном, национальном и глобальном масштабах. Этот документ является первым из серии учебных материалов, охватывающих базовые, промежуточные и продвинутые уровни в области NIR и MIR спектроскопии почв.





# 1 | Введение

Спектроскопия в видимом и ближнем инфракрасном (БИК; visible and near infrared: vis-NIR) и среднем инфракрасном (mid-infrared; MIR) диапазонах появилась и развивалась в последние три десятилетия как важный метод количественного лабораторного анализа почв (Baumgardner *et al.*, 1986; Chang *et al.*, 2001; Reeves, 2010; Viscarra Rossel *et al.*, 2006). Многие исследователи считают, что vis-NIR и MIR могут стать альтернативой традиционным лабораторным методам анализа почв, основанного на так называемой «мокрой химии» (Janik, Merry and Skjemstad, 1998; Nocita *et al.*, 2015). Для решения многих современных практических задач требуется привлечение больших объемов количественных данных о почве как в пространстве, так и во времени. Одним из примеров является точное земледелие, когда с полей регулярно отбирают пробы почв (например, по сетке) и анализируют их в лаборатории для создания картограмм почвенных свойств. Эти картограммы затем становятся основой для создания зон управления угодьями или для определения норм внесения удобрений, извести и поливных вод (Nawar *et al.*, 2017). Другим примером является анализ закрепления углерода в почве и расчет углеродных кредитов, когда на одном и том же поле регулярно отбирают пробы почв и оценивают изменения запасов органического углерода (Smith *et al.*, 2020). Огромное количество образцов почвы, которые необходимо проанализировать, требует внедрения быстрых и недорогих методов, таких как vis-NIR и MIR спектроскопия, чтобы сделать эти задачи экономически реализуемыми. Это обуславливает широкий интерес к внедрению vis-NIR и MIR для рутинного анализа почв.

## 2 | Основы использования vis-NIR и MIR излучения для анализа почв

### 2.1. Спектр электромагнитного излучения, длина волны и волновое число

Электромагнитный (ЭМ) спектр состоит из гамма-лучей, рентгеновских лучей, ультрафиолетовых, видимых, инфракрасных (ближних, средних и дальних) волн, микроволн и радиоволн, охватывающих много порядков по длине волны  $\lambda$  (или, в обратной зависимости, по частоте  $\nu$ ). Диапазоны ЭМ, главным образом используемые для анализа почвы - это vis-NIR и MIR. Длина волны vis-NIR обычно указывается в нм (нанометр,  $10^{-9}$  м) или мкм (микрометр,  $10^{-6}$  м). Vis-NIR объединяет видимую и ближнюю инфракрасную области и обычно относится к диапазону длин волн от 350 до 2500 нм (от 0,35 до 2,5 мкм). MIR обычно указывают в волновых числах ( $\text{см}^{-1}$ ), что буквально означает, сколько электромагнитных волн укладывается в одном сантиметре. MIR обычно начинается с 4000  $\text{см}^{-1}$  и заканчивается на 600 (или 400)  $\text{см}^{-1}$ , в зависимости от используемых приборов. Длина волны  $\lambda$  в нм и волновое число в  $\text{см}^{-1}$  обратно пропорциональны:

$$\text{волновое число} = \frac{10^7}{\lambda}$$

Возьмем в качестве примера длину волны 2500 нм; эквивалентное ей волновое число равно

$$\frac{10^7}{2500} = 4\,000 \text{ см}^{-1} \text{ и из этого примера видно, что конец vis-NIR совпадает с началом MIR.}$$

### 2.2. Спектры отражения и поглощения

Когда электромагнитная энергия направляется на поверхность почвы, она может поглощаться, передаваться и отражаться. Поглощенная энергия иногда может переизлучаться в виде флуоресценции.

В спектроскопии почвы наибольший интерес представляет отраженная энергия от поверхности почвы. Существует два типа отражения: зеркальное и диффузное (рассеянное). При диффузном отражении падающий световой луч отражается (рассеивается) во всех направлениях. При анализе почвы на спектрометрах используют именно режим диффузного отражения. Он предпочтителен, поскольку при диффузном отражении электромагнитная энергия проникает и в достаточной степени взаимодействует с почвенной матрицей и, следовательно, содержит больше информации о компонентах почвы. По этой причине в литературе по спектроскопии почвы часто используют такие термины, как спектроскопия диффузного отражения (Diffuse Reflectance Spectroscopy, DRS) или ИК-спектроскопия диффузного отражения с Фурье-преобразованием (Diffuse Reflectance Infrared Fourier Transform Spectroscopy, DRIFT). Также при анализе почвы применяется спектроскопия нарушенного полного внутреннего отражения (НПВО; attenuated total reflectance, ATR). Спектры почвы могут быть представлены либо спектрами отражения, либо спектрами поглощения, поэтому практикующие специалисты должны понимать, работают ли они со спектрами отражения или поглощения. Коэффициент отражения (R) и коэффициент поглощения (A) могут быть преобразованы друг в друга с помощью уравнений:  $A = \log_{10} \left( \frac{1}{R} \right)$  и  $R = 10^{-A}$ . На рис. 1 показан пример спектров отражения и поглощения vis-NIR образца почвы.

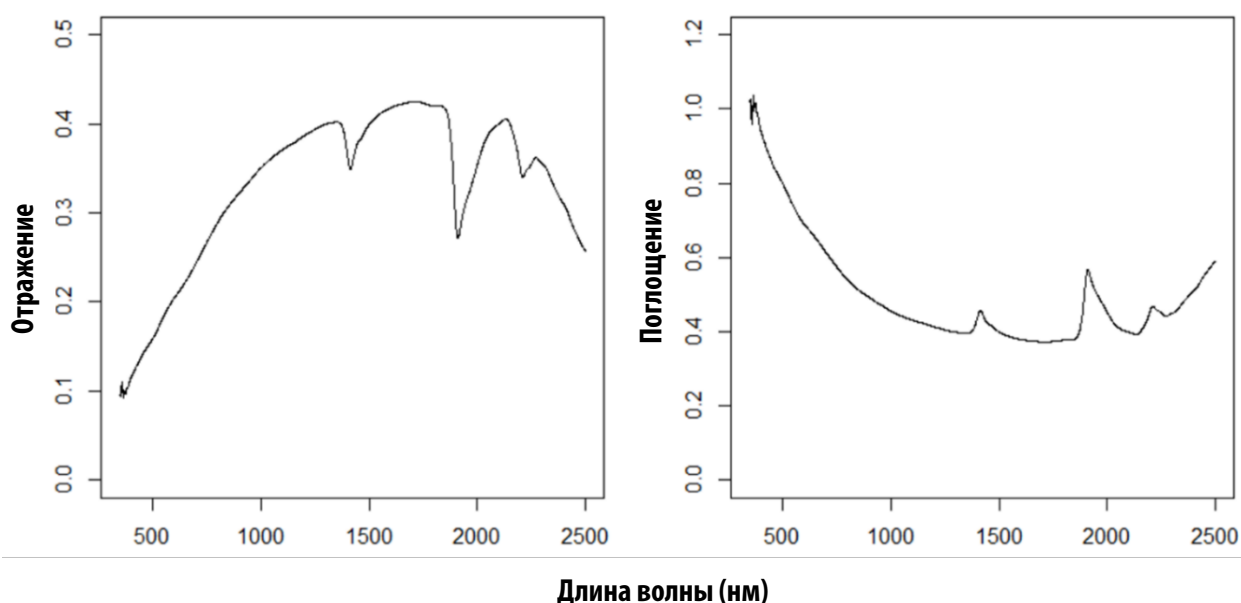


Рисунок 1. Спектры почвы в видимом и БИК-диапазонах (vis-NIR): спектр отражения (слева) и поглощения (справа)

### 2.3. Основные полосы поглощения в MIR, обертоны и комбинации в NIR

Еще одно важное различие, которое необходимо провести между vis-NIR и MIR излучением – это виды взаимодействия между электромагнитной энергией и почвой. Энергия MIR излучения вызывает колебания химических связей, обычно присутствующих как в органических, так и в минеральных соединениях почв. В органических соединениях (т.е. органическом веществе) соответствующими химическими связями являются O-H, N-H, C-H, C-C, C=C, C-N, C=O и т.д. В минеральных соединениях важны химические связи Al-OH и Si-OH, присутствующие в глинах. Эти основные колебательные полосы поглощения обычно являются сильными и хорошо разрешенными, поэтому модели MIR превосходят модели vis-NIR в прогнозировании свойств почвы, таких как содержание и состав органического вещества и глинистых минералов. В области NIR появляются обертоны и комбинации этих основных полос. Например, характерные особенности поглощения NIR-спектров почвы при длине волны около 1450 и 1920 нм связаны с обертонами колебаний растяжения O-H и H-O-H свободной воды. Полосы поглощения в NIR слабее, сильнее перекрываются и менее различимы, чем в MIR. Кроме того, в видимом диапазоне присутствие оксидов двухвалентного

и трехвалентного железа вызывает особенности поглощения из-за электронных переходов катионов железа. Наконец, более высокое содержание органического вещества в почве имеет тенденцию снижать коэффициент отражения в видимом диапазоне, придавая почве более темный цвет.

#### **2.4. Свойства почвы, которые могут быть прямо или косвенно оценены с помощью спектров в vis-NIR и MIR диапазонах**

Почва представляет собой сложную смесь из широкого набора химических компонентов и имеет различные физические состояния с точки зрения размеров частиц, агрегации, шероховатости поверхности и содержания воды. Некоторые из этих физических и химических составляющих взаимодействуют с энергией vis-NIR и MIR излучения и создают особенности поглощения в спектрах, что, по сути, является основой спектроскопии почвы с помощью vis-NIR и MIR излучения. Эти “первичные” свойства почвы, включая органическое вещество (или органический углерод), карбонаты (или неорганический углерод), общий азот, глинистые минералы, содержание железа, фракции частиц по размерам глины, ила и песка и содержание воды, обычно можно откалибровать по спектрам почвы в vis-NIR и MIR диапазонах, поскольку полосы поглощения в спектрах соответствуют этим минеральным и органическим почвенным компонентам. Важно отметить, что с помощью vis-NIR и MIR излучения также можно оценивать те свойства почвы, которые напрямую не взаимодействуют с энергией vis-NIR или MIR излучения. Например, почвенные катионы, такие как Mg или Ca, не вызывают активного спектрального поглощения, однако их часто можно достаточно хорошо оценить с помощью vis-NIR и MIR излучения, главным образом благодаря вторичной корреляции с глинистыми минералами и почвенными карбонатами. Аналогичным образом, такие свойства почвы как pH, ЕКО (катионообменная способность), засоленность и содержание питательных веществ (например, общего фосфора и калия), также могут быть оценены посредством их корреляции с одним или несколькими спектрально активными “первичными” свойствами почвы.

В литературе есть учебники и ссылки, в которых обобщены основные полосы поглощения компонентов почвы в MIR диапазоне, их обертоны и комбинационные полосы в NIR диапазоне, а также электронные переходы в видимом диапазоне. Заинтересованные читатели для ознакомления могут обратиться к *Viscarra Rossel et al., 2016* и *Soriano-Disla et al., 2014*. Важно отметить, что, несмотря на то, что многие компоненты почвы и химические группы могут быть отнесены к характеристичным полосам поглощения в спектрах vis-NIR и MIR диапазона, эти полосы поглощения редко используются отдельно для оценки свойств почвы. Основными подходами в почвенной спектроскопии в vis-NIR и MIR диапазонах скорее являются многовариантное моделирование и методы машинного обучения, охватывающие все диапазоны длин волн. В дополнение к “первичным” и “вторичным” проблемам корреляции, упомянутым выше, также важно понимать, что подходы к моделированию, основанные на спектрах в vis-NIR и MIR диапазонах, являются эмпирическими и в значительной степени управляемыми большими данными. Эти подходы (которые будут дополнительно рассмотрены в разделе 3.5) часто приводят к получению сложных численных моделей, которые нелегко интерпретировать. Эти модели подвержены чрезмерной аппроксимации, и всегда есть вероятность получить точные результаты, основанные на ложных корреляциях, обнаруженных в спектрах. Поэтому практикующим специалистам рекомендуется воздерживаться от стремления использовать модели, которые, как кажется, лучше всего подходят для имеющихся данных, игнорируя при этом другие знания о почве, имеющие отношение к анализу; например, рассмотрение вопроса о том, действительно ли оцениваемое свойство почвы прямо или косвенно спектрально активно vis-NIR и MIR в диапазонах.

В то время как анализ и моделирование отдельно спектров почв в vis-NIR или MIR диапазонах более распространены, объединение спектров vis-NIR и MIR для оценки свойств почвы может использовать преимущества обеих спектральных областей и повысить точность оценки. Для этого лаборатория должна

быть оснащена приборами, осуществляющими спектроскопический анализ в обоих диапазонах, что часто требует от лаборатории значительных первоначальных финансовых вложений.

## 2.5. Преимущества анализа почв с применением vis-NIR и MIR спектроскопии

Существует несколько явных преимуществ vis-NIR и MIR спектроскопии для анализа почвы по сравнению с традиционными методами анализа почвы, основанными на «мокрой химии». Во-первых, получение спектральных данных в vis-NIR и MIR диапазонах является быстрым и неразрушающим; получение одного сканирования занимает всего несколько секунд, и в процессе сканирования не расходуется почвенный материал. Во-вторых, спектр от одного сканирования позволяет одновременно оценивать несколько свойств почвы на одном и том же объеме почвенного материала. Это отличается от традиционных лабораторных методов, где каждое свойство почвы анализируется независимо на разных подвыборках. Таким образом, при использовании почвенной спектроскопии удается избежать потенциального варьирования, возникающего из-за варьирования внутри подвыборок. В-третьих, после первоначальной стоимости приобретения спектрометра это недорогая и экологически чистая технология, поскольку она требует минимальной подготовки образца (в основном сушки и измельчения) и не требует химических реагентов. Хотя спектрометры в vis-NIR и MIR диапазонах могут быть дорогими. Например, стоимость спектрометра ASD Labspec для анализа почвы в vis-NIR диапазоне колеблется от 50 000 до 65 000 долл. США; а прибор Bruker FTIR с высокопроизводительным оборудованием для отбора проб для анализа в MIR диапазоне стоит более 100 000 долл. США; эти цены могут также существенно различаться от страны к стране и от региона к региону. Но соотношение первоначальных инвестиций в оборудование и его отдачи является высоким, поскольку сканируются и анализируются десятки тысяч образцов почвы, что дает большой выигрыш по сравнению с первоначальными инвестициями. В совокупности эти преимущества приводят к количественному анализу почвы с гораздо более высокой производительностью и меньшими затратами, чем при использовании обычных лабораторных методов «мокрой химии».

Стоит отметить, что совсем недавно на рынке появились более дешевые портативные NIR-спектрометры с ограниченным диапазоном длин волн. Однако не все спектрометры могут дать точную оценку свойств почвы, обычно это зависит от охвата длин волн. Возможности этих ограниченных по длинам волн NIR-спектрометров для анализа почв были изучены в нескольких исследованиях (Shariffar *et al.*, 2019; Tang, Jones and Minasny, 2020).

## 3 | Процедуры анализа почв с применением vis-NIR и MIR спектроскопии

Основной рабочий процесс анализа почвы на основе vis-NIR и MIR спектроскопии представлен на рисунке 2. Vis-NIR и MIR спектроскопия отличается от других традиционных методов анализа почвы тем, что vis-NIR или MIR спектрометры напрямую не измеряют свойства почвы количественно. Вместо этого используется набор данных для обучения (также известный как калибровочный набор), чтобы обучить (или откалибровать) эмпирические модели, которые связывают спектральные данные с целевыми свойствами почвы. Поэтому образцы почвы в обучающем наборе должны быть проанализированы с помощью эталонного лабораторного аналитического метода (традиционный анализ методами «мокрой химии») в качестве общепринятого стандарта. Эмпирические модели необходимо протестировать на «независимом» наборе выборок, не используемом при калибровке модели (тестовый набор, который также анализируется тем же эталонным методом) для оценки предсказательной силы модели. После полной калибровки и тестирования модели затем могут быть использованы для оценки свойств новых,

неизвестных образцов почвы по их спектрам в vis-NIR или MIR диапазонах. Обучение и тестирование моделей являются основой анализа в vis-NIR и MIR диапазонах и требуют определенного набора знаний и навыков (например, обработка большого объема спектральных данных, статистическое моделирование, программирование на компьютерном языке и оценка предсказательной силы модели) по сравнению с обычным лабораторным анализом почвы. Этапы анализа спектральных данных в vis-NIR и MIR диапазонах описаны в следующих подразделах.

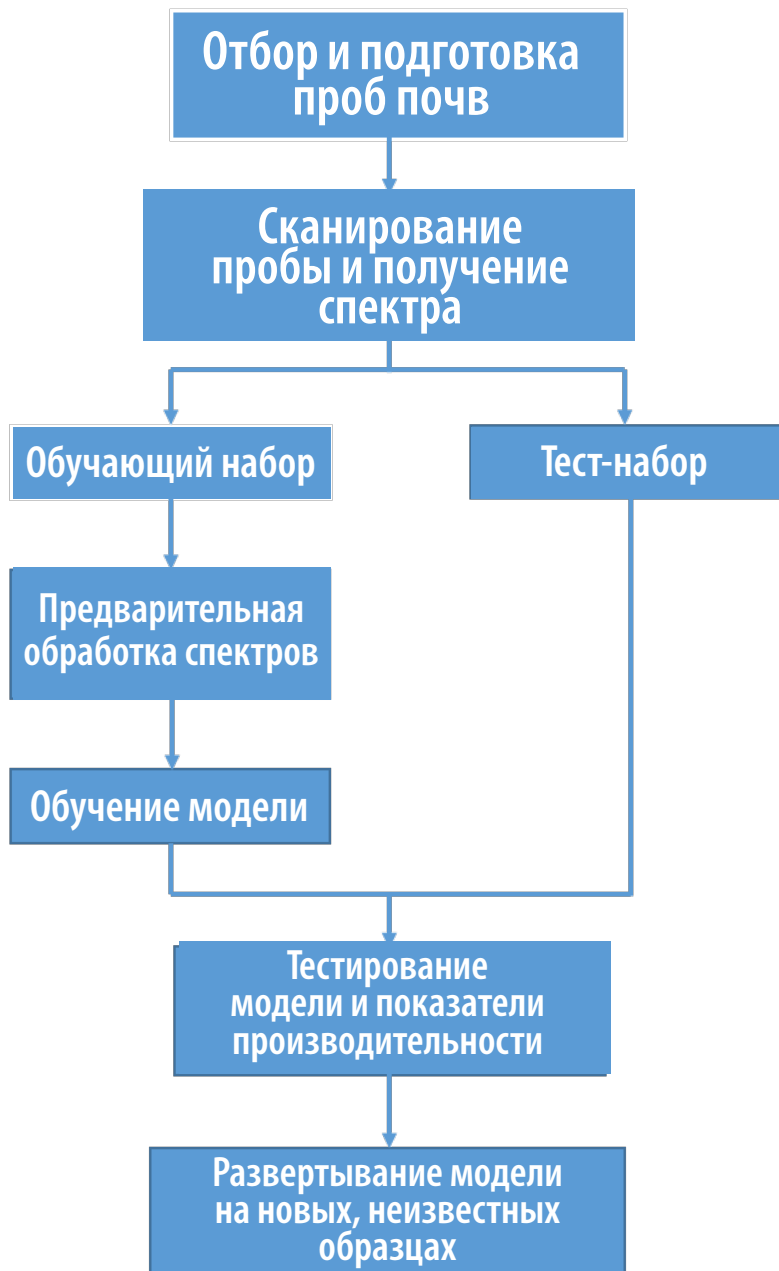


Рисунок 2. Рабочий процесс vis-NIR и MIR спектроскопии для анализа почв

### 3.1. Подготовка образцов

При полевом проботборе почв необходимо удалять крупные включения и корни растений. Образцы почвы высушивают на воздухе (35-40° С) до постоянной массы и измельчают, чтобы пропустить через 2-миллиметровое сито. Это единственная пробоподготовка, необходимая перед vis-NIR-сканированием. Необходимо обеспечить, чтобы новые образцы находились в тех же условиях, что и образцы, использованные ранее при лабораторном анализе. Для анализа vis-NIR обычно не проводят измельчение

образцов до более мелких размеров частиц, поскольку размеры частиц влияют на проникновение и рассеяние энергии vis-NIR и тонкое измельчение сделает результаты менее сопоставимыми с другими исследованиями (Nduwamungu *et al.*, 2009). Для MIR-сканирования образцы должны быть дополнительно тонко измельчены до размера частиц менее 100 мкм. В литературе есть исследования, показывающие, что тонкое измельчение значительно улучшает производительность анализа MIR (Wijewardane *et al.*, 2021). Тонкое измельчение является дополнительным этапом подготовки образца и увеличивает время и стоимость MIR-сканирования (но считается необходимым для обеспечения наилучшего качества MIR-сканирования). Таким образом, лаборатория должна инвестировать в механическую мельницу для тонкого измельчения образцов почвы. Сам процесс тонкого измельчения занимает значительно больше времени, включая время, необходимое для загрузки и выгрузки каждого образца в измельчитель, а также для тщательной очистки и сушки держателей образцов, чтобы избежать перекрестного загрязнения образцов почвы. Поэтому тонкое измельчение считается фактором, ограничивающим скорость анализа в MIR диапазоне. В этом видео, подготовленном Министерством сельского хозяйства США и Службой охраны природных ресурсов (USDA-NRCS), подробно описана процедура тонкого измельчения почвы: <https://www.youtube.com/watch?v=6RpK3zUWMaQ>.

### 3.2. Сканирование спектров

Для получения спектра диффузного отражения образцов почвы в vis-NIR диапазоне используется сертифицированная стандартная панель с коэффициентом отражения более 99 процентов на всех длинах волн (фактически считается идеальным 100-процентным диффузным отражателем). Этот процесс называется выравниванием баланса белого. В MIR диапазоне в качестве стандарта отражения используется шероховатая металлическая поверхность (например, алюминий). Спектрометр регистрирует три измерения:  $DN_{White}$  в качестве исходного спектра для стандарта отражения ( $DN$  означает дискретное число),  $DN_{Dark}$  в качестве исходного спектра для темнового тока (тока, возникающего без воздействия света) и  $DN_{Soil}$  в качестве исходного спектра для образца почвы. Коэффициент отражения почвы рассчитывается как 
$$= \frac{DN_{Soil} - DN_{Dark}}{DN_{White} - DN_{Dark}}$$
; и это преобразование обычно выполняется прибором автоматически. Во время сеанса следует периодически проводить измерения эталонного белого и темнового тока, чтобы убедиться, что прибор хорошо откалиброван. В лаборатории, где такие факторы окружающей среды, как температура и влажность, стабильны с течением времени, калибровку прибора можно проводить каждые 15 минут. Однако в самых сложных и требовательных анализах этот этап калибровки выполняется между каждым образцом.

Еще одним важным шагом для обеспечения высококачественных спектров почвы является включение прибора на достаточно длительное время (в зависимости от типа и марки прибора) перед проведением каких-либо измерений. Спектральный выход источников света в vis-NIR и MIR диапазонах, а также чувствительность спектральных детекторов зависят от температуры. Фактически, многие приборы vis-NIR и MIR спектроскопии используют определенные типы охлаждения (например, либо термоэлектрическими средствами, либо жидким азотом), чтобы снизить тепловой шум в детекторах. Этот шаг предназначен для того, чтобы избежать негативного влияния первоначального прогрева прибора на спектральные данные.

Протоколы и принадлежности, с помощью которых образцы почв помещаются в приборы различных марок, различаются. Здесь мы приводим описание аксессуаров для спектрометров ASD, которые наиболее широко используются для почвенной спектроскопии. ASD имеет два типа насадок: «кружка» (muglight) и контактный датчик. Образцы почв (воздушно-сухие, < 2 мм) загружают в держатель-шайбу, а затем помещают в «кружку» (muglight) для сканирования. Боросиликатные чашки Петри (такие как Duroplan) также часто используются для хранения образцов почвы и сканирования с помощью mug-

light. Контактный датчик является портативным, более гибким, чем muglight, и чаще используется для сканирования образцов в пакете или почвы в естественном состоянии. Тщательная очистка держателей образцов (шайб или чашек Петри) и кварцевого окна на датчике muglight/contact между образцами важна для предотвращения перекрестного загрязнения.

Спектральные данные почвы сначала сохраняются в приборе в фирменных форматах. Эти приборы также предоставляют пакеты программного обеспечения для базовой обработки данных, очистки и спектрального моделирования, такие как программное обеспечение Indigo Pro от ASD и программное обеспечение OPUS от Bruker. Тем не менее, почвенное сообщество переходит на пакеты программного обеспечения с открытым исходным кодом, такие как язык программирования R (<https://www.r-project.org>) или Python для обработки и моделирования спектральных данных почвы. Эти программные пакеты предоставляют мощные функции и библиотеки; и они особенно эффективны при обработке больших наборов данных (например, содержащих тысячи выборок) или с использованием передовых алгоритмов машинного обучения для спектрального моделирования. Необработанные спектральные данные преобразуются в файлы TXT (с разделителями табуляции) или CSV (переменные, разделенные запятыми), которые содержат матрицу спектральных данных, прежде чем они будут считаны в R или Python. В этих файлах образцы почвы упорядочены по строкам, а длины волн (или волновое число) упорядочены по столбцам.

В следующих разделах представлены этапы предварительной обработки спектра, обучения и тестирования модели, методы спектрального моделирования и оценки предсказательной силы моделей. Для демонстрации результатов этих шагов подготовлены примеры наборов данных в vis-NIR и MIR диапазонах. Набор данных в vis-NIR диапазоне содержит спектральные данные из 201 образца почвы, измеренные спектрометрией ASD LabSpec. Набор данных в MIR диапазоне содержит спектральные данные 540 образцов, измеренных спектрометром Bruker Vertex 70 FT-IR. Лабораторные данные, использованные для демонстрации спектрального моделирования, включают содержание органического углерода (измеренного сухим сжиганием), ила (измеренное методом пипетки) и pH (экстракция водой 1:1). Оба набора данных извлекаются из открытых спектральных библиотек в vis-NIR и MIR диапазонах, хранящихся в Лаборатории почвенных исследований Келлога Министерства сельского хозяйства США (Kellogg Soil Survey Lab of USDA-NRCS). Кроме того, предоставлены R-коды для иллюстрации концепций, демонстрации этапов моделирования и воспроизведения результатов ([EduSoils - Soil spectral data analysis \(Level I\) - Introduction #1](#)).

### **3.3. Предварительная обработка спектров**

Этап предварительной обработки спектров обычно применяется к необработанным спектральным данным в vis-NIR и MIR диапазонах. Предварительная обработка спектров может уменьшить случайный шум в необработанных спектрах, улучшить отношение сигнал/шум, свести к минимуму посторонний эффект рассеяния света, уменьшить размерность спектральных данных для эффективного вычисления и/или улучшить характеристики поглощения. Наиболее простым методом предварительной обработки является скользящее среднее/бинирование, которое просто усредняет спектральные данные по длине волны/волновому числу (например, каждые 10 или 20 точек данных) для одновременного уменьшения шума и размерности. Другими обычно используемыми методами предварительной обработки при анализе почвы являются: (1) стандартное отклонение случайной величины с нормальным распределением (standard normal variate SNV), (2) мультипликативная коррекция сигнала (multiplicative signal correction MSC), (3) фильтрация Савицкого и Голея (SG) и (4) удаление континуума. Программа R имеет специальный пакет под названием `prospectr` (<https://CRAN.R-project.org/package=prospectr>) для предварительной спектральной обработки.

SNV нормализует каждую строку спектральной матрицы путем вычитания каждой строки из ее среднего значения и деления ее на стандартное отклонение. SNV - это простой способ нормализации спектральных данных, предназначенный для коррекции рассеяния света.

$$SNV_i = \frac{x_i - \bar{x}_i}{s_i}$$
 где  $\bar{x}_i$  - i-й необработанный спектр,  $\bar{x}_i$  - его среднее значение, а  $s_i$  - стандартное отклонение i-го спектра.

MSC первоначально означает "мультипликативная корреляция рассеяния", но значение изменилось с годами, потому что оно также полезно для других типов мультипликативных задач, помимо рассеяния. Основная концепция MSC заключается в устранении нелинейностей в данных, вызванных рассеянием частиц в образцах. MSC реализуется путем выравнивания каждого необработанного спектра  $\bar{x}_i$  с идеальным эталонным спектром  $\bar{x}_r$ , в предположении, что  $\bar{x}_i = m_i \bar{x}_r + a_i$ , следующим образом:

$$MSC_i = \frac{x_i - a_i}{m_i}$$
 где  $a_i$  и  $m_i$  - аддитивные и мультипликативные члены, соответственно. При реализации MSC идеальный эталонный спектр  $\bar{x}_r$  обычно недоступен. Для этой цели часто используется средний спектр набора данных.

Фильтрация SG (Savitzky and Golay, 1964) широко используется для предварительной обработки спектров почв в vis-NIR и MIR диапазонах. Этот метод универсален и может использоваться для сглаживания/шумоподавления и дифференциации. В фильтрации SG есть три параметра. Первый параметр - это размер окна (2g+1). Второй параметр определяет порядок применения дифференцирования (например, p = 0 означает сглаживание; p = 1 означает первую производную; p = 2 означает вторую производную). Третий параметр - это порядок полиномов, который определяет степень сглаживания. Одним из недостатков фильтрации SG является то, что при размере окна 2g+1 результирующий спектр потеряет начальную и конечную точки данных g. Например, если используется окно из 11 точек (g = 5), результирующий спектр будет на 5 точек короче как в начале, так и в конце.

Наконец, удаление континуума, предпринимаемое для удаления непрерывной части спектра, а также выделения особенностей поглощения, полезно при визуализации тонких, перекрывающихся полос поглощения в vis-NIR диапазоне. Однако использовать для моделирования спектр, удаленный от континуума, не принято.

Выбор методов предварительной обработки несколько субъективен и зависит от конкретного случая. Многие исследователи предпочитают использовать одновременно несколько методов предварительной обработки (например, мультипликативную коррекцию сигнала с последующей фильтрацией Савицкого-Голея), в то время как другие исследователи обнаружили, что предварительная обработка не обязательно улучшает предсказательную силу модели. Общего руководства по использованию методов предварительной обработки спектра не существует, и пользователям рекомендуется использовать эвристический подход, попробовать несколько методов предварительной обработки и выбрать те, которые существенно улучшают производительность модели. На рисунке 3 показаны результаты применения этих методов предварительной обработки спектра.



### Длина волны (нм)

Примеры спектров отражения в vis-NIR диапазоне из экспериментального набора данных (n=201) с различными методами предварительной обработки спектра: (А) исходные спектры отражения; (Б) спектры поглощения; (С) спектры после стандартного нормального изменения SNV; (D) спектры после мультипликативной коррекции сигнала MSC; (Е) спектры после удаления континуума; и (F) спектры первой производной после фильтрации Савицкого-Голея. Синие линии - средние спектры, рассчитанные по всем образцам.

### 3.4. Обучение и тестирование модели

Как указывалось ранее, весь экспериментальный набор данных случайным образом разбивается на два набора, один для обучения модели, а другой как “независимый” набор для тестирования модели. Обычное разделение - это 70 процентов для обучения и 30 процентов для тестирования. Разделение набора данных на обучающий и тестовый наборы называется разделением данных. В идеале процесс случайного разделения данных повторяется несколько раз, чтобы гарантировать низкую случайную вероятность получения хорошего / плохого прогноза. Другими методами являются К-кратная перекрестная проверка (cross-validation CV) и бутстреппинг. В CV набор данных разбивается на К непересекающихся подмножеств. Подмножества К-1 используются для обучения модели, в то время как предсказательная сила модели оценивается на оставшемся подмножестве. Обычно К устанавливается равным 10 (так называемое 10-кратное CV), но оно может быть равно количеству спектров; как в случае CV без учета одного. При бутстреппинге выбирается несколько подвыборок с заменой для подгонки и оценки модели. Оба метода являются лучшей альтернативой оценке модели, чем разделение данных, поскольку статистика для оценки модели не чувствительна к единичному рандомному разделению данных, которое происходит случайно.

Целью обучения модели является получение многомерной эмпирической модели (например, линейной регрессии), которая связывает спектральные данные с целевыми свойствами почвы. Поскольку спектральные данные обычно содержат сотни или тысячи сильно коллинеарных полос, достижение баланса между “предсказуемостью” и “переобучением” является ключом к обучению модели. Переобучение относится к ситуации, когда модель демонстрирует отличную предсказательную силу в обучающем наборе, но плохо работает в независимом тестовом наборе; другими словами, модель не обладает способностью к обобщению. Внутренняя перекрестная проверка обучающих данных — это метод, используемый для поиска “наилучшей модели”, которая не перегружает обучающий набор, используя такую статистику, как внутренняя среднеквадратичная ошибка CV (Root Mean Squared Error of CV,  $RMSE_{cv}$ ), которую можно сравнить со сложностью модели для достижения баланса между ними.

Пример оценки переобучения модели для частичной регрессии наименьших квадратов (Partial Least Squares Regression, PLSR) показан на рисунке 4, где “обучение” относится к  $RMSE$  повторной замены, а “CV” относится к значению  $RMSE$ , полученному из 10-кратного CV. В PLSR количество скрытых переменных (latent variables LV) является параметром модели, указывающим на сложность модели. По мере того, как в модель включается все больше LVs,  $RMSE$  повторного представления продолжает уменьшаться. Однако для 10-кратного CV значение  $RMSE$  достигает минимума, когда количество LVs равно 13. Следовательно, мы приходим к выводу, что модели PLSR, содержащие более 13 LVs, перегружены, на что указывает их более низкий  $RMSE$  обучения, но более высокий  $RMSE$  перекрестной проверки.

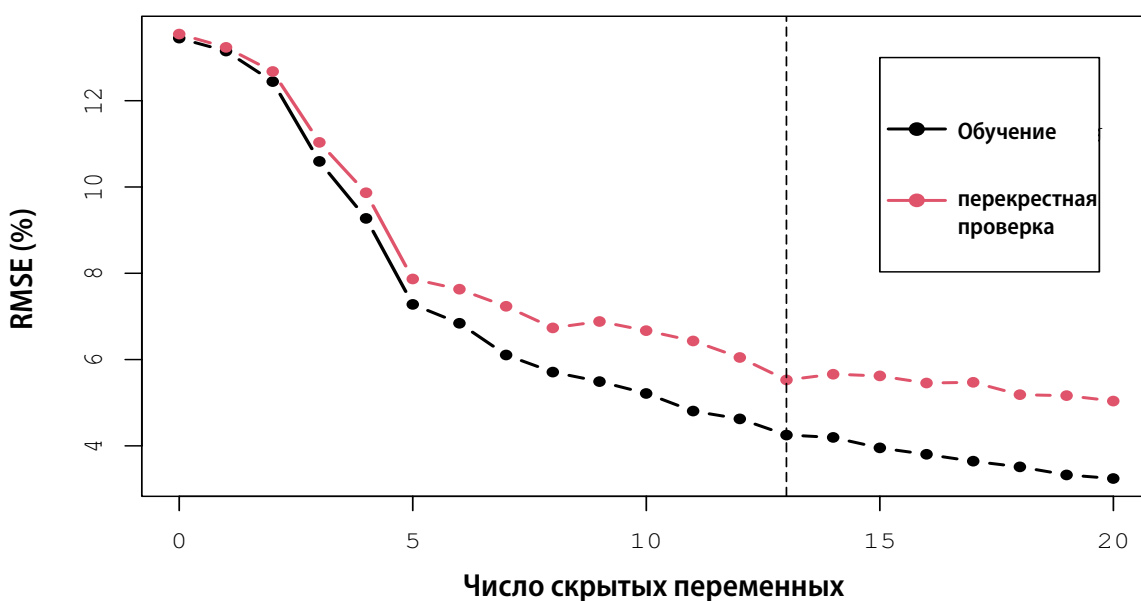


Рисунок 4. Иллюстрация переобучения модели с частичной регрессией наименьших квадратов

### 3.5. Методы обучения модели

Хотя спектральные полосы поглощения, связанные с определенными компонентами почвы, идентифицированы и обобщены в литературе, сами по себе эти полосы редко используются для построения количественных моделей для прогнозирования содержания этих компонентов почвы с помощью хорошо известного закона Ламберта-Бэра. Вместо этого основные подходы заключаются в обучении эмпирических многомерных моделей, которые связывают спектральные данные в vis-NIR и MIR диапазонах с целевыми свойствами почвы. В ранних исследованиях для выбора подмножества

спектральных полос для обучения модели использовалась множественная линейная регрессия (multiple linear regression, MLR) (Ben-Dor and Banin, 1995). Позже приобрели популярность и широко использовались в спектроскопии почв методы моделирования, включая регрессию частных наименьших квадратов (Partial Least Squares Regression, PLSR) и регрессию главных компонент (Principal Component Regression, PCR) (см., например, Chang *et al.*, 2001 и Viscarra Rossel *et al.*, 2006). PLSR стал де-факто стандартным методом, с которым сравниваются другие новые методы моделирования.

Основные различия между PLSR/PCR и множественной линейной регрессией заключаются в следующем: (1) PLSR и PCR используют весь спектр для моделирования, тогда как MLR выбирает только небольшое подмножество диапазонов волн для моделирования; и (2) PLSR и PCR используют матричное проектирование для проецирования исходных спектральных данных и формирования набор новых переменных, известных как скрытые переменные (LV) или главные компоненты (PC). Эти новые переменные не коррелируют друг с другом, и на первые несколько LV или PC приходится большая доля общей дисперсии в исходных спектральных данных. Эти два свойства, LV и PC, эффективно решают проблему мультиколлинеарности спектральных данных в vis-NIR/MIR диапазонах, а также проблему большого  $p$  и малого  $n$  для небольших наборов выборок. Здесь  $p$  относится к числу переменных (спектральных полос), а  $n$  относится к числу выборок при моделировании. Подробная математика PLSR и PCR выходит за рамки этого вводного документа. Заинтересованные читатели могут обратиться к Wold, Sjöström и Eriksson, 2001 для описания принципа и реализации PLSR и PCR.

Совсем недавно получило широкое распространение в vis-NIR и MIR спектроскопии большое семейство так называемых методов “машинного обучения”. Примерами таких методов являются алгоритм случайного леса (Random Forest, RF), регрессия опорных векторов (Support Vector Regression, SVM), кубист, искусственные нейронные сети (Artificial Neural Networks, ANN) и т.д. (Minasny and McBratney, 2008; Viscarra Rossel *et al.*, 2016). Модели, обученные с помощью этих методов машинного обучения, обычно демонстрируют более высокую точность прогнозирования, чем PLSR и PCR, особенно при большом количестве обучающих выборок. Одним из объяснений превосходной предсказательной силы этих моделей является то, что они могут более эффективно моделировать нелинейные взаимосвязи между спектральными данными и свойствами почвы. Методы машинного обучения иногда критикуют за их природу “черного ящика”, что приводит к плохой интерпретируемости модели. Например, очень трудно визуализировать, интерпретировать или объяснить модель SVM или ANN, в терминах которой полосы поглощения играют более важную роль в прогнозировании определенных свойств почвы. Для реализации этих методов также может потребоваться больше вычислительных ресурсов. Таким образом, во многих современных приложениях обучение модели реализуется на графических процессорах (графический процессор) или кластере графических процессоров / процессоров.

Метод главных компонент (МГК) (Principal Component Analysis, PCA) — это концепция, тесно связанная с PCR и PLSR. Фактически, PCR - это множественная линейная регрессия значений главных компонент (principal component, PC), полученных в результате МГК. В спектроскопии почвы МГК обычно выполняется в качестве первого шага для изучения внутренней изменчивости спектральных данных в vis-NIR или MIR диапазонах и выявления спектральных выбросов. Обратите внимание, что в исходном спектральном пространстве довольно сложно быстро идентифицировать выбросы из группы спектров (как показано на рис. 3). На рисунке 5А показаны спектральные данные в MIR диапазоне для 540 образцов. Очевидно, что при таком представлении трудно определить, какие выборки являются выбросами. На рисунке 5В показан график осыпи первых 10 главных компонент, описывающий процент отклонения, приходящийся на каждую компоненту (на первую компоненту приходится 78 процентов от общей дисперсии, на вторую - 8,6 процента, на третью - 4,2 процента и так далее). Часто нас интересует, на сколько компонент в совокупности приходится более 90%, 95% или 99% общей дисперсии спектральных данных. В этом примере на первые

три компонента приходится >90 процентов, а на первые 5 - более 95 процентов общей дисперсии. На рисунке 5С показана попарная диаграмма рассеяния выборки, установленной в пространстве компонент (PC1 против PC2 против PC3), каждая точка представляет одну пробу и соответствует одной кривой в исходном спектральном пространстве на рисунке 5А. Синий круг - это 99-процентный доверительный интервал. Точки, лежащие снаружи, являются потенциальными спектральными выбросами, которые заслуживают дальнейшего изучения.

Поглощение

волновое число, см<sup>-1</sup>

Дисперсия

Главные компоненты

**Рисунок 5. Иллюстрация МГК (PCA) на спектральных данных почвенной MIR-спектроскопии: (А) данные почв в MIR диапазоне в спектральном пространстве; (В) график осыпи первых 10 основных компонент и процентная дисперсия, приходящаяся на каждую компоненту; и (С) попарные диаграммы рассеяния первых трех компонент**

### 3.6. Оценка модели

Для прогнозирования свойств новых, неизвестных образцов почвы необходима объективная оценка эффективности моделей в vis-NIR и MIR диапазонах. Для этой цели используется независимый тестовый набор, в котором свойства почв измеряются с помощью стандартных методов. Предсказательная сила моделей обычно оценивается путем сравнения прогнозов в vis-NIR или MIR диапазонах с эталонными значениями и вычисления статистических данных, таких как смещение (средняя ошибка, mean error, ME), среднеквадратичная ошибка (root mean squared error, RMSE), коэффициент корреляции Пирсона r и коэффициент детерминации (R<sup>2</sup>).

Средняя ошибка:

$$ME = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)$$

где  $\bar{y}$  и  $\hat{\bar{y}}$  обозначают измеренные и прогнозируемые значения свойств почвы с помощью модели в vis-NIR или MIR диапазонах, соответственно, и  $\sqrt{N}$  - общее количество измеренных значений.

Среднеквадратичная ошибка:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}$$

Коэффициент корреляции Пирсона:

$$r = \frac{\sum_{i=1}^N (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^N (y_i - \bar{y})^2} \sqrt{\sum_{i=1}^N (\hat{y}_i - \bar{\hat{y}})^2}}$$

Где  $\bar{y}$  - среднее значение измеренных значений, а  $\bar{\hat{y}}$  - среднее значение прогнозируемых значений.

ME представляет смещение, а RMSE - величину ошибки. RMSE - это показатель того, насколько прогноз близок к эталонному значению для отдельного образца, и, следовательно, показатель ожидаемой точности прогноза, когда модель в vis-NIR или MIR диапазонах применяется к неизвестному образцу почвы. Смещение, с другой стороны, является мерой тенденции модели переоценивать или недооценивать определенные свойства почвы. Оба имеют оптимальное значение 0, но ME может быть отрицательной. Коэффициент корреляции  $r$  является мерой линейной корреляции между прогнозируемыми и эталонными значениями и находится в диапазоне от -1 (идеальная отрицательная линейная корреляция) до 1 (идеальная положительная линейная корреляция). Одним из неудобств использования RMSE в качестве показателя предсказательной силы модели является то, что он зависит от свойств и единиц измерения. Это затрудняет сравнение моделей в vis-NIR или MIR диапазонах с различными свойствами почвы и различными исследованиями. Например, модель pH почвы в vis-NIR диапазоне имеет RMSE 0,45 единицы pH; модель органического углерода почвы в vis-NIR диапазоне может иметь RMSE 0,35 процента; тогда как модель в vis-NIR диапазоне для содержания ила имеет RMSE в 5 процентов. Без дополнительной информации трудно определить, какая модель в vis-NIR диапазоне работает лучше всего.

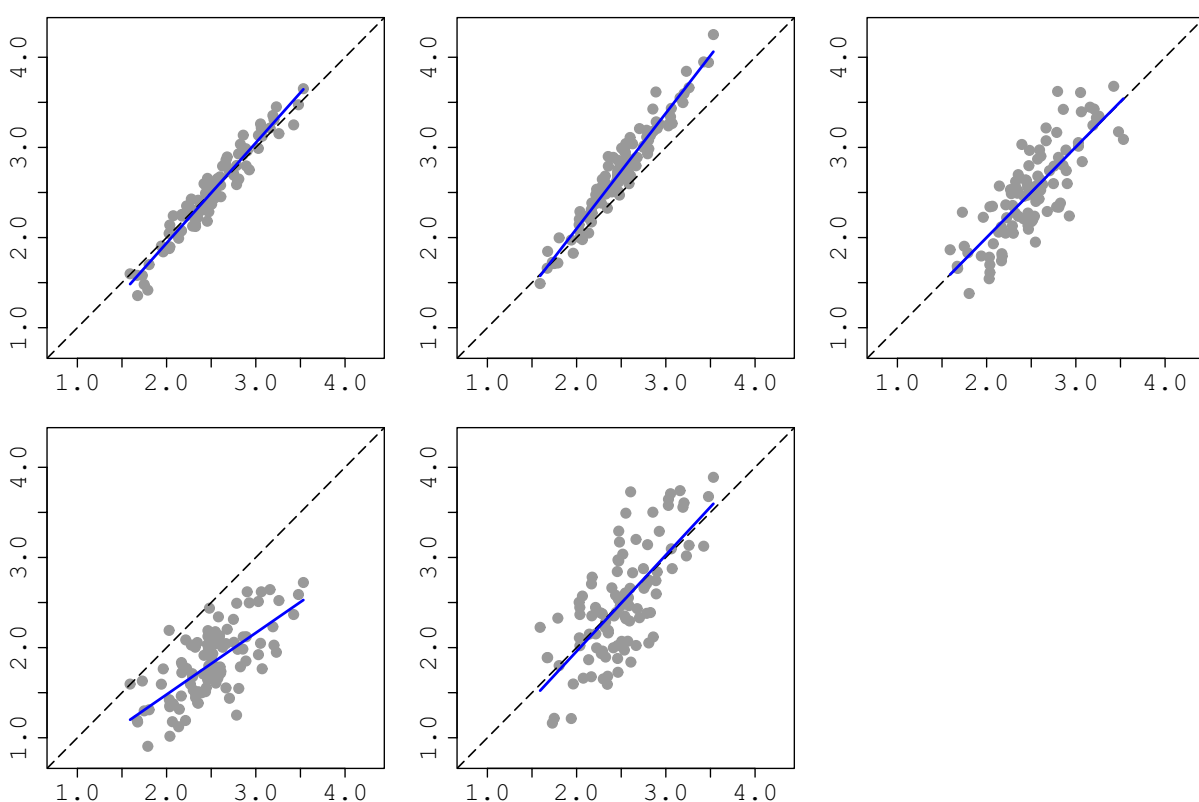
Наконец, коэффициент детерминации оценивается по формуле:

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2}$$

что эквивалентно  $1 - \text{RMSE}^2 / \sigma_y^2$ , где  $\sigma_y^2$  отклонение измеренных значений. В литературе  $R^2$  также упоминается как эффективность моделирования или коэффициент Нэша-Сатклиффа. Оптимальное значение  $R^2$  равно 1, и оно может быть отрицательным. Когда  $R^2 < 0$ , это указывает на то, что среднее значение наблюдений является лучшим предиктором, чем модель. Положительное значение может быть интерпретировано как величина отклонения, объясняемая моделью. Например, модель pH почвы в vis-NIR диапазоне с  $R^2 = 0,3$  означает, что 30 процентов отклонения измеренных значений pH приходится на модель в vis-NIR диапазоне. Обратите внимание, что  $R^2$  часто вычисляется как квадрат коэффициента корреляции Пирсона  $r$  (т.е.  $r^2$ ).  $R^2$  и  $r^2$  эквивалентны только в случае линейной модели с свободным коэффициентом.  $R^2$  оценивает отклонение от линии 1:1, тогда как  $r^2$  оценивает отклонение от установленной линии линейной регрессии между измеренным и прогнозируемым. Мы подчеркиваем, что для реалистичной оценки характеристик спектроскопической модели коэффициент детерминации должен быть рассчитан как  $1 - \text{RMSE}^2 / \sigma_y^2$ .

Мы также отмечаем, что некоторые публикации оценивают модель на основе отношения производительности к отклонению (performance to deviation, RPD), которое представляет собой индекс производительности модели, аналогичный  $R^2$ . В качестве альтернативы было предложено использовать отношение производительности модели к межквартильному расстоянию (performance to interquartile distance, RPIQ) для учета ненормально распределенных данных путем замены стандартного отклонения в уравнении RPD на межквартильный диапазон данных. При оценке модели не существует золотого стандарта, исходя из которого модель считается достаточной, удобство использования и качество модели зависят от ее применения.

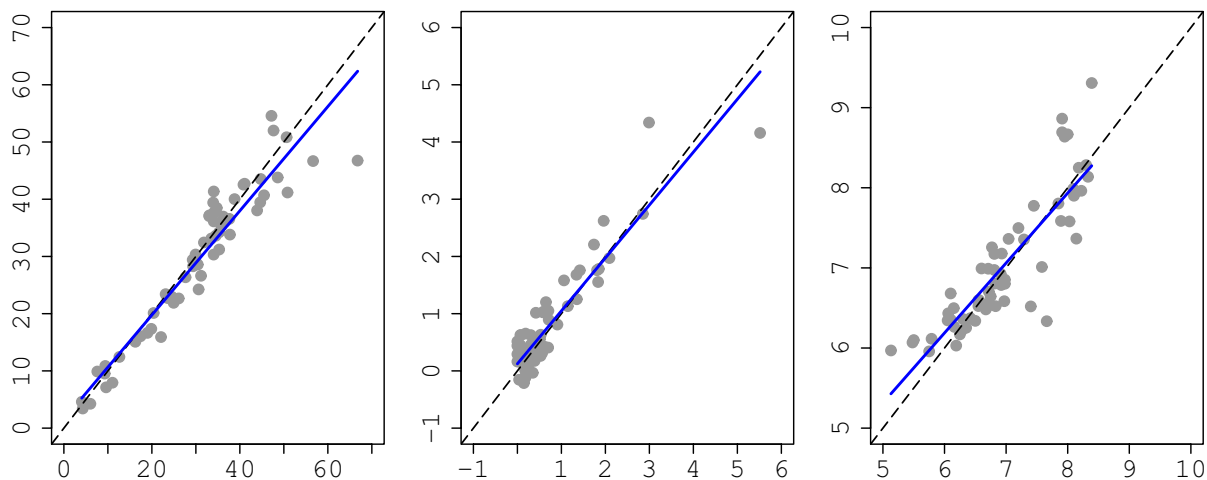
На рисунке 6 использованы смоделированные данные для иллюстрации того, насколько прогнозируемые значения в vis-NIR / MIR диапазонах сравнимы с измеренными в лаборатории значениями для определенных значений ME, RMSE,  $r$  и  $R^2$ . Обратите внимание на плотное прилегание точек вокруг линии регрессии и на то, как линии регрессии отклоняются от линии 1:1 (при оценке по  $R^2$ ).



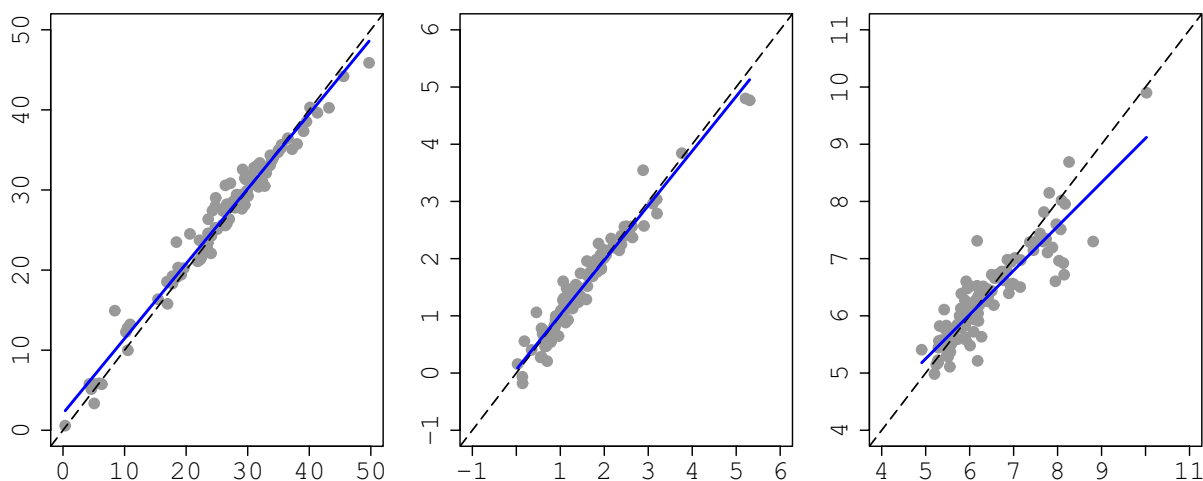
**Рисунок 6. Смоделированные данные, показывающие сравнение предсказанных значений в vis-NIR или MIR диапазоне свойств почвы с измеренными в лаборатории свойствами на диаграммах рассеяния. Приведены различные уровни производительности модели и приведены показатели их оценки.**

Ниже приведен пример использования наборов данных в vis-NIR и MIR диапазонах из Лаборатории почвенных исследований Келлога Службы охраны природных ресурсов Министерства сельского хозяйства США (USDA-NRCS) для оценки трех свойств почвы: органического углерода, содержания ила и pH. Набор данных в vis-NIR диапазоне содержит в общей сложности 201 выборку, и он был разделен на 70 процентов для обучения ( $n=141$ ) и 30 процентов для тестирования ( $n=60$ ). Набор данных в MIR диапазоне содержит 540 выборок, и он был разделен на 80% для обучения ( $n=432$ ) и 20% для тестирования ( $n=108$ ). Для моделирования использовался метод PLSR. Результаты тестирования представлены в виде диаграмм рассеяния на рис. 7 (для vis-NIR) и рис. 8 (для MIR); а числовые значения ME, RMSE,  $r$  и  $R^2$  приведены в

таблице 1. Кроме того, в таблице 1 также перечислены характеристики моделей регрессии опорных векторов (Support Vector Regression, SVR) для этих двух наборов данных. SVR - это распространенная модель машинного обучения в исследованиях почвенной спектроскопии, и мы использовали ее здесь в качестве примера. Обратите внимание, что на рис. 7, рис. 8 и в таблице 1 показаны общие способы представления результатов моделирования в vis-NIR/MIR диапазонах.



**Рисунок 7.** Диаграммы рассеяния предсказанных значений в vis-NIR диапазоне и измеренных в лаборатории свойств почвы для тестового набора (n = 60) с использованием частичной регрессии наименьших квадратов. Пунктирная серая линия - линия равенства между наблюдаемым и прогнозируемым, синяя линия – линия тренда



**Рисунок 8.** Диаграммы рассеяния предсказанных MIR и измеренных в лаборатории свойств почвы для тестового набора (n = 108) с использованием частичной регрессии наименьших квадратов

Таблица 1. Результаты тестирования трех почвенных свойств с моделированием PLSR и SVR для значений в vis-NIR и MIR диапазонах

Моделирование	Почвенное свойство	vis-NIR (n = 60)				MIR (n = 108)			
		ME	RMSE	r	R <sup>2</sup>	ME	RMSE	r	R <sup>2</sup>
PLSR	Ил	-1.16	4.35	0.95	0.9	0.41	1.7	0.98	0.97
	C орг	0.07	0.38	0.92	0.83	-0.01	0.2	0.97	0.95
	pH	0.07	0.42	0.87	0.72	-0.09	0.43	0.89	0.79
SVR	Ил	-1.68	5.53	0.92	0.83	0.39	2.35	0.97	0.94
	C орг	-0.1	0.44	0.92	0.77	0.05	0.2	0.98	0.95
	pH	-0.08	0.49	0.8	0.63	-0.02	0.44	0.88	0.77

## 4 | Спектральные библиотеки почв в vis-NIR и MIR диапазонах в региональном, континентальном и глобальном масштабах: актуальность, преимущества и ограничения

Наиболее дорогостоящей частью спектроскопии в vis-NIR и MIR диапазонах для анализа почвы является разработка модельного обучающего набора данных, который необходимо не только сканировать приборами в vis-NIR и MIR диапазонах, но и анализировать с помощью эталонных лабораторных методов для определения свойств почвы. На этапе исследований общий подход заключается в разделении всего набора данных на обучающий и тестовый наборы для обучения и тестирования моделей, соответственно. Однако для разработки обучающего набора для каждого проекта такой подход не является ни практичным, ни экономичным. Более жизнеспособным подходом была бы разработка библиотек почвенных спектральных данных, которые могли бы использоваться совместно многими пользователями и проектами. Это, по сути, мотивирует разработку спектральных почвенных библиотек в vis-NIR и MIR диапазонах в региональном, национальном, континентальном и глобальном масштабах. Существует несколько способов создания таких спектральных библиотек. Во-первых, старые образцы почвы могут быть извлечены из почвенных архивов и отсканированы с помощью прибора в vis-NIR и / или MIR диапазонах. Поскольку старые образцы почвы уже высушены, измельчены и для них имеются определенные химические данные, это представляет собой более дешевый способ создания спектральной библиотеки. Некоторый недостаток старых образцов почвы заключается в том, что не все представляющие интерес свойства почвы будут в наличии или проанализированы с помощью желательных эталонных методов. В этом случае можно провести повторный анализ некоторых устаревших образцов. В качестве альтернативы для создания спектральной библиотеки почв могут быть проведены кампании по отбору проб новых почвенных образцов более систематическим и контролируемым образом. Спектральные почвенные библиотеки в vis-NIR и MIR диапазонах часто содержат большое количество образцов почвы (например, десятки тысяч), охватывают большие географические регионы и, вероятно, объединяют образцы из разных кампаний и проектов по отбору проб. Эти факторы создают дополнительные проблемы, связанные с разработкой и использованием спектральных библиотек. Во-первых, одни и те же свойства почвы могут быть измерены различными лабораторными методами. Например, в качестве эталонных методов определения органического углерода в почве используются как "сухое сжигание", так и метод Уокли-Блэйка. Содержание органического углерода в почве, определенное по Уокли-Блэйку, может быть использовано для оценки органического вещества почвы с использованием общего коэффициента



пересчета 1,724. Но метод Уокли-Блэйка основан на окислении с использованием раствора  $K_2Cr_2O_7$  и может окислять не весь органический углерод почвы. Следовательно, истинный коэффициент пересчета может зависеть от конкретной почвы. С другой стороны, при сухом сжигании измеряются все формы углерода (включая неорганический углерод карбонатов). рН почвы можно измерить в воде, растворе  $KCl$  или  $CaCl_2$  при различных соотношениях почва: раствор. Для определения содержания питательных веществ почвы, таких как фосфор и калий, используются различные методы экстракции. Если образцы почвы, измеренные разными эталонными методами или из разных лабораторий с разными протоколами, будут внесены в одну и ту же спектральную библиотеку, база данных будет противоречивой. При неправильном обращении это несоответствие будет усилено во время спектроскопического моделирования и приведет к снижению предсказательной силы модели.

Вторая проблема крупномасштабных спектральных библиотек заключается в чрезвычайно большой вариабельности образцов почвы (материнских пород, минералогического состава, физических и химических свойств). Эта вариабельность также приводит к большой изменчивости спектральных данных в *vis-NIR* и *MIR* диапазонах. С другой стороны, при решении практических задач необходимость оценки свойств почвы часто возникает в местном масштабе (например, в точном земледелии для оценки плодородия почвы и его изменения в полевых условиях). Чтобы более эффективно использовать спектральные библиотеки, необходимо выбрать и оптимизировать подмножество выборок для калибровки модели. Для решения этой проблемы разрабатываются более продвинутое методы машинного обучения (например, поиск спектрального сходства и обучение на основе памяти) и схемы моделирования (например, добавление дополнительного веса) для повышения производительности прогнозирования свойств почвы с использованием спектров, которые спектрально похожи, вместо использования всех спектральных данных в библиотеке. Также обычно проверяют, действительно ли новый образец, свойства которого нужно прогнозировать, попадает в пространство, охватываемое спектрами в библиотеке. Если определено, что новый образец не попадает в спектральное пространство библиотеки, то этот образец нельзя оценивать при помощи библиотеки, а необходимо провести традиционный лабораторный анализ. Такую оценку можно провести с помощью метода главных компонент и путем рисования выпуклого контура вокруг пространства, охватываемого спектрами библиотеки. Один из таких методов был недавно описан в (Wadoux *et al.*, 2021).

Третья проблема связана с вариабельностью, вносимой в крупномасштабные спектральные библиотеки в результате использования различных приборов, сканирующих в *vis-NIR* и *MIR* диапазонах. Лаборатории, которые вносят свой вклад в наполнение библиотеки, вероятно, будут использовать приборы разных производителей или разные модели от одного и того же производителя приборов. Протоколы анализов и способы сканирования, а также лабораторная среда (температура, влажность и вентиляция), в которой расположены приборы, могут привести к значительным различиям в спектральных данных (Ge, Thomasson and Sui, 2011). Разработка стандартных операционных протоколов для работы прибора и сбора спектральных данных для совместного использования лабораториями поможет согласовать спектральные данные из различных источников. Для решения этой проблемы также следует рассмотреть возможность пересчета калибровки (Feudale *et al.*, 2002), который позволяет моделям, откалиброванным на одном приборе, быть полезными для спектральных данных с другого прибора. Несмотря на упомянутые выше предостережения, эти спектральные библиотеки представляют собой важный шаг вперед в получении недорогих количественных данных о почве для таких задач, как точное земледелие и секвестрация углерода в почве. Например, фермерам необходимо разработать карты плодородия почвы и рН, чтобы грамотно проводить внесение удобрений и известки с переменной нормой внесения путем интенсивного сбора образцов почвы (например, один образец на акр или 2,5 образца на гектар). Для большого поля общее количество образцов почвы составило бы пару сотен. С помощью спектральных библиотек в *vis-NIR* и *MIR* диапазонах образцы нужно только сканировать с помощью *vis-NIR* или *MIR* спектрометра, а такие

свойства почвы, как органическое вещество и pH, можно оценить с помощью существующих моделей. Это значительно экономит время и затраты, связанные с подготовкой и анализом образцов. В таблице 2 мы обобщили опубликованные спектральные почвенные библиотеки в vis-NIR и MIR диапазонах в региональном, континентальном и глобальном масштабах.

**Таблица 2. Краткое описание опубликованных почвенных спектральных библиотек в vis-NIR и MIR диапазонах в национальном, континентальном и глобальном масштабах**

Авторы и год	Масштаб	Количество образцов	Представленные лабораторные данные	Примечание
<b>Vis-NIR</b>				
Brown <i>et al.</i> , 2006	Глобальный	3768 из США и 416 со всего мира	Глина, Сорг, С неорг, Fe, ЕКО, глинистые минералы	Архивные образцы с 1988 по 1999 год, охарактеризованные Национальным центром почвенных исследований США – Лабораторией почвенных исследований; ASD FieldSpec
Brodský <i>et al.</i> , 2011	Национальный (Чехия)	5 223	pH, ЕКО, TC, IC, P-Мэйлих, К, Са, Mg	Архивные образцы Кафедры почвоведения и защиты почв Чешского Университета естественных наук в Праге; ASD FieldSpec 3
Viscarra Rossel and Webster, 2012	Национальный (Австралия)	21 493	24 почвенных свойства	В основном Национальный почвенный архив CSIRO, датируемый 70 годами; ASD LabSpec
Stevens <i>et al.</i> , 2013	Континентальный (Европа)	~20 тыс.	Глина, пыль, песок, pH, ЕКО, ОС, IC, TN, P, К	23 страны ЕС, образцы, собранные в ходе обследования LUCAS в 2008-2012 годах; Поверхностные образцы (0-30 см); FOSS NIRSystems
Araújo <i>et al.</i> , 2014	Национальный (Бразилия)	7 172	ОВ, глина, пыль, песок	Университет Сан-Паулу; Образцы профиля; ASD Field Spec Pro FR
Shi <i>et al.</i> , 2014	Национальный (Китай)	1 581	ПОВ	Четырнадцать провинций Китая; Образцы поверхности (0-20 см); ASD FieldSpec Pro FR

Clairotte <i>et al.</i> , 2016	Национальный (Франция)	~3 800	ОС	Образцы из сетки участков на территории французской метрополии, две глубины (0-30, 30-50 см); ASD LabSpec 2500
Wijewardane <i>et al.</i> , 2016	Национальный (США)	~20 тыс.	ОС, ТС	Проект быстрой оценки выбросов углерода US-DA-NRCS; Профильные образцы; ASD LabSpec
Viscarra Rossel <i>et al.</i> , 2016	Глобальный	23 631	ОС, IC, pH, CEC, Fe, глина, пыль, песок	Образцы из 92 стран мира; Серия ASD (FieldSpec, AgriSpec, TerraSpec, LabSpec)
<b>MIR</b>				
Terhoeven-Urselmans <i>et al.</i> , 2010	Глобальный	971	pH, ОС, ЕКО, Mg, глина, песок, Ca	Образцы из 18 стран, Международный справочно-информационный центр почв; Профильные образцы; Bruker Tensor 27 FTIR
Viscarra Rossel <i>et al.</i> , 2008	Национальный (Австралия)	1878 (213 с лабораторными данными)	pH, УЭП, ЕКО, ОС, K, Na, Mg, Ca, глина, пыль, песок	Устаревшие образцы, четыре основных региона выращивания хлопка; Профильные образцы; Bruker Tensor 37 FTIR
Grinand <i>et al.</i> , 2012	Национальный (Франция)	~2 000	ОС, IC	Образцы, собранные в 2002-2009 годах; Поверхностные образцы (0-30 см); Thermo Nicolet 6700
Clairotte <i>et al.</i> , 2016	Национальный (Франция)	~3 800	ОС	Образцы из сетки участков на территории метрополии Франции; Две глубины (0-30, 30-50 см); Thermo Nicolet 6700
Wijewardane <i>et al.</i> , 2018	Национальный (США)	~20 000	ОС, IC, ТС, TN, TS, глина, пыль, песок, ЕКО, K, P, pH	Министерство сельского хозяйства США -Национальный центр изучения почв NRCS, образцы, собранные с 2001 по 2018 год; Профильные образцы; Bruker Vertex 70 FTIR

## 5 | Приборы для сканирования почвы в vis-NIR и MIR диапазонах

Хотя в принципе можно использовать любые спектрометры с аксессуарами для рассеянного отражения, сообщество почвоведов выделяет некоторые предпочтительные марки приборов для сканирования почвы. Что касается vis-NIR диапазона, то спектрорадиометры серии ASD (FieldSpec и Lab-Spec), по-видимому, чаще всего используются при анализе почв. ASD была ранее известна как Analytical Spectral Devices (базируется в Боулдере, штат Колорадо, США) и в настоящее время является частью Malvern Panalytical. Другие основные приборы для vis-NIR сканирования включают FOSS, SpectralEvolution и Thermo Nicolet (Benedetti and van Egmond, 2021).

Для почвенной MIR-спектроскопии часто используются FTIR-системы производства Bruker Optics (Alpha, HTS-XT) и Thermo Nicolet (6700). Компания Agilent выпустила на рынок портативные приборы (4100 Exoscan и 4300 Handheld), которые могут проводить измерения в полевых условиях (Benedetti and van Egmond, 2021).

Подробный список распространенных приборов для анализа почв в vis-NIR и MIR диапазонах можно найти в недавно опубликованной глобальной оценке по почвенной спектроскопии, проведенной Глобальным почвенным партнерством ФАО (Benedetti and van Egmond, 2021). Как описано выше, появляются недорогие портативные NIR-спектрометры, которые работают в определенном диапазоне длин волн. Было показано, что такие устройства, как сканер Neospectra (Si-Ware), FT-NIR с длиной волны 1250-2500 нм, являются хорошей альтернативой спектрометрам исследовательского класса.

## 6 | Заключительные замечания

Спектроскопия почв в видимом и ближнем инфракрасном (БИК; visible and near infrared: vis-NIR) и среднем инфракрасном (mid-infrared; MIR) диапазонах уже давно предложена и исследована в качестве быстрого, недорогого и надежного метода количественного лабораторного анализа почв. Многочисленные тематические исследования по всему миру показали, что с помощью vis-NIR и MIR спектроскопии можно точно оценивать множество физических и химических свойств почвы. По сравнению с традиционными аналитическими методами, vis-NIR и MIR спектроскопия требует только сушки на воздухе и измельчения образцов и, следовательно, является несложным методом с точки зрения навыков, необходимых для подготовки образцов. Эксплуатация приборов, измеряющих спектры в vis-NIR и MIR диапазонах, осуществляется в соответствии с набором установленных (и стандартизированных) процедур и согласованных усилий для обеспечения высококачественных спектральных измерений. Спектральные данные содержат тысячи точек данных, на основе которых необходимо разработать статистические модели для оценки свойств почвы. Это самое важное различие между vis-NIR/MIR спектроскопией и другими лабораторными аналитическими методами. По этой причине персонал лаборатории должен обладать различным набором навыков в области интерпретации результатов, проверки ошибок и контроля/обеспечения качества. С созданием больших спектральных библиотек в vis-NIR и MIR диапазонах в локальном, региональном и континентальном масштабах становится важной пространственно распределенная глобальная сеть лабораторий почвенных анализов, сосредоточенная исключительно на vis-NIR и MIR спектроскопии. Совместное использование спектральных библиотек и моделей среди сети почвенных лабораторий значительно повысит скорость и снизит стоимость анализа почвы, и в то же время сделает результаты более сопоставимыми и воспроизводимыми среди лабораторий сети.

# Библиография

- Araújo, S.R., Wetterlind, J., Demattê, J.A.M. & Stenberg, B.** 2014. Improving the prediction performance of a large tropical vis-NIR spectroscopic soil library from Brazil by clustering into smaller subsets or use of data mining calibration techniques. *European Journal of Soil Science*, 65(5): 718–729. <https://doi.org/10.1111/ejss.12165>
- Baumgardner, M.F., Silva, L.F., Biehl, L.L. & Stoner, E.R.** 1986. Reflectance Properties of Soils. *Advances in Agronomy*, pp. 1–44. Elsevier. [https://doi.org/10.1016/S0065-2113\(08\)60672-0](https://doi.org/10.1016/S0065-2113(08)60672-0)
- Ben-Dor, E. & Banin, A.** 1995. Near-Infrared Analysis as a Rapid Method to Simultaneously Evaluate Several Soil Properties. *Soil Science Society of America Journal*, 59(2): 364–372. <https://doi.org/10.2136/sssaj1995.03615995005900020014x>
- Benedetti, F. & van Egmond, F.** 2021. *Global Soil Spectroscopy Assessment*. FAO. <https://doi.org/10.4060/cb6265en>
- Brodský, L., Klement, A., Penížek, V., Kodešová, R. & Borůvka, L.** 2011. Building soil spectral library of the Czech soils for quantitative digital soil mapping. *Soil and Water Research*, 6(4): 165–172.
- Brown, D.J., Shepherd, K.D., Walsh, M.G., Mays, M.D. & Reinsch, T.G.** 2006. Global soil characterization with VNIR diffuse reflectance spectroscopy. *Geoderma*, 132(3–4): 273–290. <https://doi.org/10.1016/j.geoderma.2005.04.025>
- Chang, C.W., Laird, D.A., Mausbach, M.J. & Hurburgh, C.R.** 2001. Near-infrared reflectance spectroscopy-principal components regression analyses of soil properties. *Soil Science Society of America Journal*, 65(2): 480–490.
- Claïrotte, M., Grinand, C., Kouakoua, E., Thébault, A., Saby, N.P.A., Bernoux, M. & Barthès, B.G.** 2016. National calibration of soil organic carbon concentration using diffuse infrared reflectance spectroscopy. *Geoderma*, 276: 41–52. <https://doi.org/10.1016/j.geoderma.2016.04.021>
- Feudale, R.N., Woody, N.A., Tan, H., Myles, A.J., Brown, S.D. & Ferré, J.** 2002. Transfer of multivariate calibration models: a review. *Chemometrics and Intelligent Laboratory Systems*, 64(2): 181–192. [https://doi.org/10.1016/S0169-7439\(02\)00085-0](https://doi.org/10.1016/S0169-7439(02)00085-0)
- Ge, Y.F., Thomasson, J.A. & Sui, R.X.** 2011. Remote sensing of soil properties in precision agriculture: A review. *Frontiers of Earth Science*, 5(3): 229–238. <https://doi.org/10.1007/s11707-011-0175-0>
- Grinand, C., Barthes, B.G., Brunet, D., Kouakoua, E., Arrouays, D., Jolivet, C., Caria, G. et al.** 2012. Prediction of soil organic and inorganic carbon contents at a national scale (France) using mid-infrared reflectance spectroscopy (MIRS). *European Journal of Soil Science*, 63(2): 141–151. <https://doi.org/10.1111/j.1365-2389.2012.01429.x>
- Janik, L.J., Merry, R.H. & Skjemstad, J.O.** 1998. Can mid infrared diffuse reflectance analysis replace soil extractions? *Australian Journal of Experimental Agriculture*, 38(7): 681. <https://doi.org/10.1071/EA97144>
- Malone, B., Wadoux, A.M.J.-C., McBratney, A.B., Fajardo, M. & Minasny, B.** 2021. *Soil Spectral Inference with R: Analysing Digital Soil Spectra Using the R Programming Environment*. Springer Nature. 257 pp.
- Minasny, B. & McBratney, A.B.** 2008. Regression rules as a tool for predicting soil properties from infrared reflectance spectroscopy. *Chemometrics and Intelligent Laboratory Systems*, 94(1): 72–79. <https://doi.org/10.1016/j.chemolab.2008.06.003>
- Nawar, S., Corstanje, R., Halcro, G., Mulla, D. & Mouazen, A.M.** 2017. Delineation of Soil Management Zones for Variable-Rate Fertilization. *Advances in Agronomy*, pp. 175–245. Elsevier. <https://doi.org/10.1016/bs.agron.2017.01.003>
- Nduwamungu, C., Ziadi, N., Tremblay, G.F. & Parent, L.-É.** 2009. Near-Infrared Reflectance Spectroscopy Prediction of Soil Properties: Effects of Sample Cups and Preparation. *Soil Science Society of America Journal*, 73(6): 1896–1903. <https://doi.org/10.2136/sssaj2008.0213>
- Nocita, M., Stevens, A., van Wesemael, B., Aitkenhead, M., Bachmann, M., Barthès, B., Ben Dor, E. et al.** 2015. Chapter Four – Soil Spectroscopy: An Alternative to Wet Chemistry for Soil Monitoring. In L.S. Donald, ed. *Advances in Agronomy*, pp. 139–159. Academic Press.
- Reeves, J.B.** 2010. Near- versus mid-infrared diffuse reflectance spectroscopy for soil analysis emphasizing carbon and laboratory versus on-site analysis: Where are we and what needs to be done? *Geoderma*, 158(1–2): 3–14. <https://doi.org/10.1016/j.geoderma.2009.04.005>
- Rossel, R.A.V., Jeon, Y.S., Odeh, I.O.A. & McBratney, A.B.** 2008. Using a legacy soil sample to develop a mid-IR spectral library. *Soil Research*, 46(1): 1. <https://doi.org/10.1071/SR07099>
- Rossel, R.A.V. & Webster, R.** 2012. Predicting soil properties from the Australian soil visible-near infrared spectroscopic database. *European Journal of Soil Science*, 63(6): 848–860. <https://doi.org/10.1111/j.1365-2389.2012.01495.x>
- Savitzky, A. & Golay, M.J.E.** 1964. Smoothing + Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry*, 36(8): 1627–1629. <https://doi.org/10.1021/Ac60214a047>
- Shariffar, A., Singh, K., Jones, E., Ginting, F.I. & Minasny, B.** 2019. Evaluating a low-cost portable NIR spectrometer for the prediction of soil organic and total carbon using different calibration models. *Soil Use and Management*, 35(4): 607–616. <https://doi.org/10.1111/sum.12537>
- Shi, Z., Wang, Q., Peng, J., Ji, W., Liu, H., Li, X. & Viscarra Rossel, R.A.** 2014. Development of a national VNIR soil-spectral library for soil classification and prediction of organic matter concentrations. *Science China Earth Sciences*, 57(7): 1671–1680. <https://doi.org/10.1007/s11430-013-4808-x>
- Smith, P., Soussana, J., Angers, D., Schipper, L., Chenu, C., Rasse, D.P., Batjes, N.H. et al.** 2020. How to measure, report and verify soil carbon change to realize the potential of soil carbon sequestration for atmospheric greenhouse gas removal. *Global Change Biology*, 26(1): 219–241. <https://doi.org/10.1111/gcb.14815>
- Soriano-Disla, J.M., Janik, L.J., Rossel, R.A.V., Macdonald, L.M. & McLaughlin, M.J.** 2014. The Performance of Visible, Near-, and Mid-Infrared Reflectance Spectroscopy for Prediction of Soil Physical, Chemical, and Biological Properties. *Applied Spectroscopy Reviews*, 49(2): 139–186. <https://doi.org/10.1080/05704928.2013.811081>

**Stevens, A., Nocita, M., Tóth, G., Montanarella, L. & van Wesemael, B.** 2013. Prediction of Soil Organic Carbon at the European Scale by Visible and Near InfraRed Reflectance Spectroscopy. *PLoS ONE*, 8(6): e66409. <https://doi.org/10.1371/journal.pone.0066409>

**Tang, Y., Jones, E. & Minasny, B.** 2020. Evaluating low-cost portable near infrared sensors for rapid analysis of soils from South Eastern Australia. *Geoderma Regional*, 20: e00240. <https://doi.org/10.1016/j.geodrs.2019.e00240>

**Terhoeven-Urselmans, T., Vagen, T.G., Spaargaren, O. & Shephard, K.D.** 2010. Prediction of Soil Fertility Properties from a Globally Distributed Soil Mid-Infrared Spectral Library. *Soil Science Society of America Journal*, 74(5): 1792–1799. <https://doi.org/10.2136/sssaj2009.0218>

**Viscarra Rossel, R.A., Behrens, T., Ben-Dor, E., Brown, D.J., Demattê, J.A.M., Shephard, K.D., Shi, Z. et al.** 2016. A global spectral library to characterize the world's soil. *Earth-Science Reviews*, 155: 198–230. <http://dx.doi.org/10.1016/j.earscirev.2016.01.012>

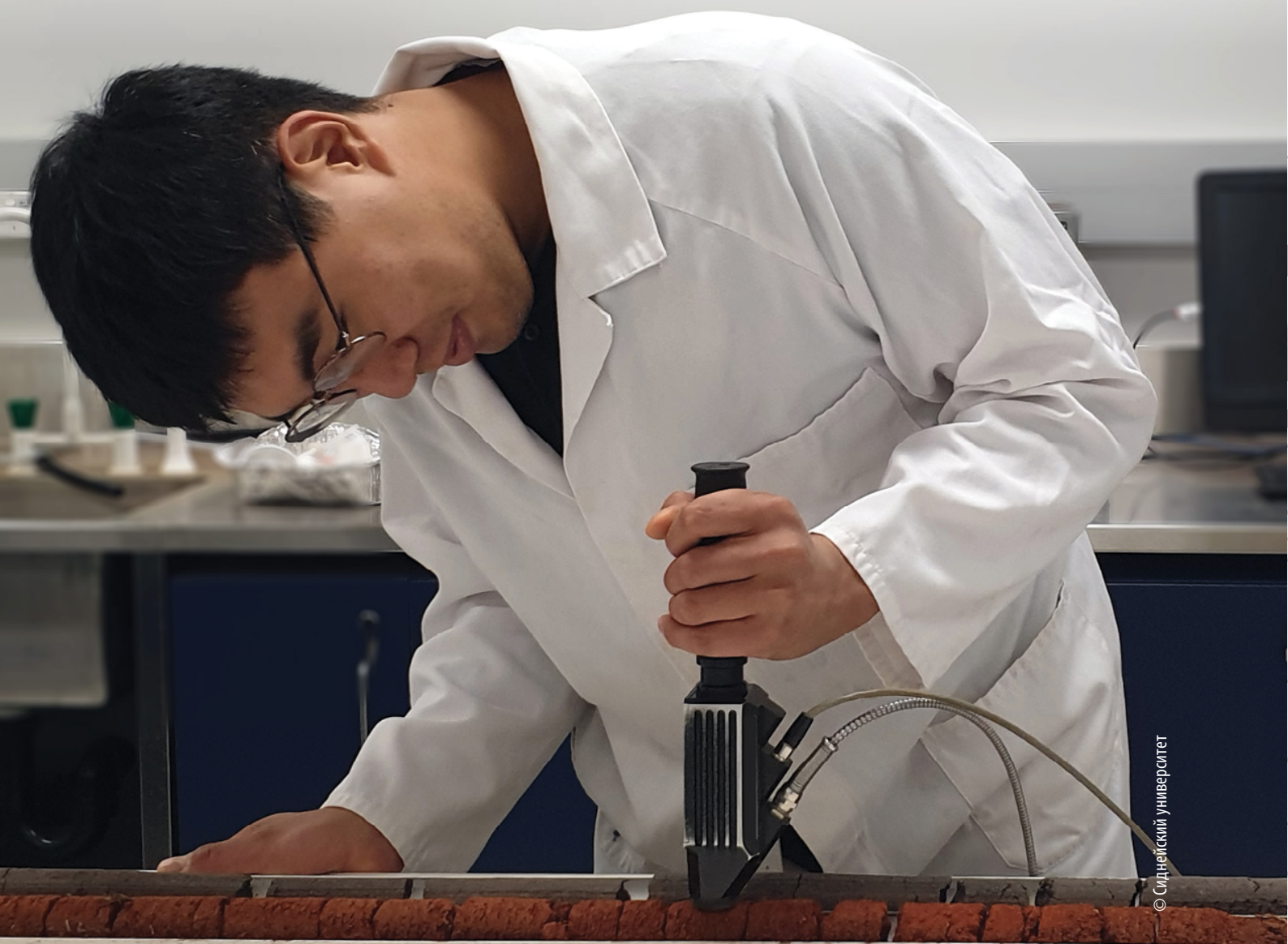
**Viscarra Rossel, R.A., Walvoort, D.J.J., McBratney, A.B., Janik, L.J. & Skjemstad, J.O.** 2006. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma*, 131(1–2): 59–75. <https://doi.org/10.1016/j.geoderma.2005.03.007>

**Wijewardane, N.K., Ge, Y., Sanderman, J. & Ferguson, R.** 2021. Fine grinding is needed to maintain the high accuracy of mid-infrared diffuse reflectance spectroscopy for soil property estimation. *Soil Science Society of America Journal*, 85(2): 263–272. <https://doi.org/10.1002/saj2.20194>

**Wijewardane, N.K., Ge, Y., Wills, S. & Libohova, Z.** 2018. Predicting physical and chemical properties of US soils with a mid-infrared reflectance spectral library. *Soil Science Society of America Journal*, 82(3): 722–731.

**Wijewardane, N.K., Ge, Y., Wills, S. & Loecke, T.** 2016. Prediction of Soil Carbon in the Conterminous United States: Visible and Near Infrared Reflectance Spectroscopy Analysis of the Rapid Carbon Assessment Project. *Soil Science Society of America Journal*, 80(4): 973–982. <https://doi.org/10.2136/sssaj2016.02.0052>

**Wold, S., Sjöström, M. & Eriksson, L.** 2001. PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, 58(2): 109–130. [http://dx.doi.org/10.1016/S0169-7439\(01\)00155-1](http://dx.doi.org/10.1016/S0169-7439(01)00155-1)







Глобальное  
почвенное  
партнерство

Глобальное почвенное партнерство (ГПП) - это глобально признанный механизм, созданный в 2012 году. Наша миссия - позиционировать почвы в Глобальной повестке дня посредством коллективных действий. Наши основные цели - продвижение устойчивого управления почвенными ресурсами и регуляторных механизмов для обеспечения здорового и продуктивного состояния почв, предоставления основных экосистемных услуг в целях продовольственной безопасности и улучшения питания, адаптации к изменению климата и смягчения его последствий, а также устойчивого развития.

### GLOSOLAN-SPEC

Инициатива Глобальной сети почвенных лабораторий по спектроскопии почв

GLOSOLAN-Spec - это глобальная инициатива по спектроскопии почвы в рамках Глобальной сети почвенных лабораторий, которая в основном направлена на развитие потенциала стран. Она включает обучение по созданию национальных/региональных почвенных спектральных лабораторий, развитие национальных/региональных почвенных спектральных библиотек со службой оценки, а также предоставление консультационных услуг по подходящему оборудованию. Цель этого развития заключается в том, чтобы предоставить странам доступ к большему количеству почвенных данных, используя эффективный по времени и затратам аналитический метод.

Благодаря финансовой поддержке



WORLD BANK GROUP



Korea  
Green Growth  
Trust Fund