



**МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М.В. ЛОМОНОСОВА**

ФАКУЛЬТЕТ ПОЧВОВЕДЕНИЯ

Утверждаю:
декан факультета почвоведения МГУ

С.А. Шоба
«___» ____ 2018 г.

**РАБОЧАЯ ПРОГРАММА ПРАКТИКИ
ВВЕДЕНИЕ В МЕТОДОЛОГИЮ АНАЛИЗА ДАННЫХ**

Направления подготовки: 06.03.02 «Почвоведение»

Авторы-составители:
д.б.н. Самсонова В.П., к.с.-х.н. Мешалкина Ю.Л.

Рабочая программа практики рассмотрена и одобрена на заседании кафедры земледелия и агроэкологии

протокол № _____ от «___» _____ 2018 г.

Заведующий кафедрой _____ (ФИО)

Рабочая программа практики утверждена на заседании учебно-методической комиссии факультета почвоведения МГУ, протокол № 2 от «17» мая 2018 г.

Председатель УМК _____ Рахлеева А.А.

Москва
2018 г.

1. Вид и тип практики, место практики в структуре образовательной программы, способ проведения:

Вид: Производственная практика

Тип: практика по получению первичных профессиональных умений и опыта профессиональной деятельности.

Вариативная часть образовательной программы.

Стационарная.

2. Входные требования для освоения практики, предварительные условия (если есть):

Высшая математика, информатика, физика, общая химия, органическая химия, аналитическая химия, физическая химия, коллоидная химия, геология с основами геоморфологии, учение о почвенных свойствах и процессах, почвоведение, химия почв, физика почв, земледелие, растениеводство

3. Результаты обучения по практике, соотнесенные с требуемыми компетенциями выпускников:

3.1. Компетенции выпускников, формируемые частично при реализации практики:

Способность осуществлять поиск, критический анализ и синтез информации (УК-1.Б)

Способность управлять своим временем, выстраивать и реализовывать траекторию саморазвития на основе принципов образования в течение всей жизни (УК-2.Б)

Способность осуществлять социальное и профессиональное взаимодействие для реализации своей роли в команде и достижения командных целей и задач (УК-12.Б)

Владение методами сбора, обработки, анализа и синтеза научной полевой и лабораторной информации в области современного теоретического, экспериментального и практического почвоведения (ОПК-1.Б).

Владение навыками культуры социальных отношений, умение работать в коллективе, способность к коллективному обсуждению проблем в области современного теоретического, экспериментального и практического почвоведения (ОПК-3.Б).

Владение системой фундаментальных научных понятий, методологией и методами современного почвоведения (ПК-1.Б)

Способность к проведению на практике производственно-изыскательских работ, почвенно-экологических исследований и исследований земельных ресурсов, с составлением отчетной документации (ПК-2.Б).

Способность эксплуатировать современную аппаратуру и оборудование для выполнения научно-исследовательских полевых и лабораторных исследований в современном теоретическом, экспериментальном и практическом почвоведении (ПК-4.Б).

3.2. Планируемые результаты обучения по практике:

В результате обучения студенты должны:

Знать организацию баз данных в пакете Excel и в среде R, библиотеки в среде R

Уметь устанавливать на персональный компьютер компоненты программной среды R и запускать Excel -online; извлекать данные по запросам и проводить их первичную статистическую обработку, пользоваться библиотеками среды R, строить разнообразные диаграммы в пакете Excel и в среде R,

Понимать процедуру первичной обработки данных в почвоведении и экологии,

Иметь практические навыки пилотного анализа данных для решения реальных ситуаций в ходе научных исследований в области почвоведения, агрохимии, экологии и природопользовании;

Иметь опыт: оценки своих ресурсов и их пределов при выполнении задания

Иметь опыт: планирования рабочего времени в полевых условиях с учетом собственных возможностей и накопленного опыта

Уметь: учитывать интересы, особенности поведения и мнения (включая критические) людей при совместной деятельности

Иметь опыт: взаимодействия с членами команды, в том числе при обмене информацией, знаниями и опытом

4. Формат обучения:

Лекции, камеральная работа

5. Объем практики составляет: 8 дней

6. Содержание практики, структурированное по темам (разделам) с указанием отведенного на них количества академических часов и виды учебных занятий

№ п/ п	Наименование разделов и тем	Трудоемкость (в академических часах) по формам занятий				Форма промежуточного контроля	
		Аудиторная работа (с разбивкой по формам и видам)					
		Лекции	Семинары	Лабораторная работа	Практические занятия		
1	Основы работы в среде Excel	2		4		4	Решение задач по теме, заполнение отчета
2	Графики и базы данных в Excel	2		6		2	Решение задач по теме, заполнение отчета
3	Основы работы в программной среде R	2		6		2	Решение задач по теме, заполнение отчета
4	Графики и базы данных в среде R	2		6		4	Решение задач по теме, заполнение отчета
5	Знакомство в пакетами в среде R, ориентированными на	2		4		2	Решение задач по теме, заполнение отчета

	научные расчеты в почвоведении и экологии					
Итого:	10		26		14	Зачет на основе составленных отчетов по задачам

Содержание практики по разделам и этапам:

Раздел 1. Основы работы в среде Excel

Этап 1. Что может быть в ячейке? (типы данных)

Как загрузить Microsoft Office online. Общая информация о программе Excel. Типы форматов: общий, числовой, денежный, финансовый, форматы дат, время, процентный, дробный, экспоненциальный и дополнительные. Уменьшение и увеличение размерности в ячейке. Примеры форматов ячеек. Разные виды форматирования ячеек. Расположение текста в ячейке. Объединение ячеек. Ввод специальных символов. Вставка примечания.

Этап 2. Расчеты по формулам

Проведение вычислений как на калькуляторе. Вычисление суммы. Важность разделителя десятичных знаков. Просмотр формул в ячейке. Индикация ошибок ввода и вычислений. Ввод формулы со ссылками на ячейки. Выявление влияющих ячеек.

Этап 3. Работа с таблицами

Адреса ячеек абсолютные и относительные. Знак \$ для фиксации ссылки на ячейки. Особенности оформления таблиц. Границы ячеек. Заливка ячеек. Выравнивание внутри ячейки. Перенос строк внутри ячеек. Пример расчета влажности почв весовым методом. Копирование автозаполнением. Копирование с помощью вставки. Особенности специальной вставки. Мастер функций. Справка по функции. Расчет сверток: среднее, минимум и максимум, число значений. Условное форматирование цветом. Получение внешних данных из интернета.

Задания для самостоятельной работы: Каждое задание выполняется студентом для своего варианта, для каждого задания выполняется отчет.

Задание 1.1 Оформление авансового отчета по результатам гипотетической поездки заграницу с пересчетом сумм в рублях и в иностранной валюте.

Задание 1.2 Вычисление потенциальной урожайности (т/га) по формуле, в зависимости от запаса продуктивной влаги в 1 м слое, ожидаемого количества осадков за вегетационный период, непроизводительного расхода почвенной влаги и урожая на единицу израсходованной воды.

Задание 1.3 Расчет весовой влажности для точек с одного поля, расположенных случайным образом. Рассчитать характеристики, используемые для свертки информации: среднее, минимум, максимум, а также число наблюдений. Провести условное форматирование значений влажности для величин, больших влажности завядания (15%). Подсчитать число значений >15, воспользовавшись функцией СЧЕТЕСЛИ(массив, условие).

Задание 1.4 Расчет индексов видового разнообразия сорной растительности для одного из участков угодья. Рассчитать следующие характеристики для варианта: индекс Шеннона, индекс Пиелу.

Раздел 2. Графики и базы данных в Excel

Этап 1. Типы исследовательских диаграмм и графиков

Диаграммы как наглядное представление информации. Гистограммы, простые графики, круговая диаграмма, график с областями, точечная диаграмма и т.п.

Этап 2. Требования к качеству диаграмм

Возможности для форматирования графиков. Основные элементы: оси, легенда, ряды данных. Макет диаграмм как способ изменить диаграмму. Форматирование рядов данных. Графики, которые не рекомендуется использовать в научных работах. Примеры неудачных график. Примеры бессмысленных диаграмм. Приемы для улучшения графиков.

Этап 3. Как организовать базу данных

Определение базы данных и ее свойства. Система управление базой данных. Типы данных. Сортировка данных: качественных и количественных. Сортировка с помощью функции Фильтр. Текстовые и числовые фильтры. Запросы в базе данных

Задания для самостоятельной работы: Каждое задание выполняется студентом для своего варианта, для каждого задания выполняется отчет.

Задание 2.1 Распределение гранулометрического состава по профилю серых лесных почв. Построить и отформатировать гистограммы распределения фракций гранулометрического состава серых лесных почв по разным глубинам (два типа гистограмм: простые и составные), построить диаграммы типа «точечная», построить круговые диаграммы для отдельных горизонтов. Всего должно быть 10 графиков или более, нужно использовать разные макеты графиков и разные цвета.

Задание 2.2 Расчет Распределение по профилю серых лесных почв подвижных форм железа и микроэлементов в виде кривых распределения по глубинам отбора образцов и в виде гистограмм по горизонтам /подгоризонтам.

Задание 2.3 Расчет сверток и построение графиков для свойств пахотного горизонта на разных элементах рельефа. Рассчитать минимум, максимум и медиану (центральное значение). Построить график, вставьте линии диапазона (Макет – Линии - Коридор диапазона) и отформатировать его.

Раздел 3. Основы работы в программной среде R

Этап 1. Общие сведения о программной среде R

Проект по разработке свободного программного обеспечения GNU. Преимущества работы с R. Формат CSV. Версии программы R. Особенности установки программ R и Rstudio. Полезные ссылки. Консоль программы R. Окна программы Rstudio. Что такое скрипт. Способы запуска команд. Вызов справки. R как калькулятор. Вызов справки.

Этап 2. Создание векторов

R – векторизованный язык. Знаки, обозначающие присвоение. Логические операторы. Как можно создать вектор. Операции с векторами. Правила векторной алгебры. Сравнение векторов. Добавление элемента в конец вектора. Перечень элементов в векторе как вектор. Задание имен элементам вектора.

Этап 3. Простейшие графики

Построение гистограмм с помощью функции barplot. Выполнение серии команд. Сохранение рисунков. Панель визуализации рисунков. Простейшие графики с помощью функции plot. Задание заголовка графика. Подписи осей. Символ точек. Тип линий.

Задания для самостоятельной работы: Каждое задание выполняется студентом для своего варианта, для каждого задания выполняется отчет.

Задание 3.1 Для 2 строк из авансового отчета (задание 1.1) рассчитать сумму в рублях, исходя из курса валюты и затрат. Прокомментировать результаты. Для двух образцов почв из своего варианта рассчитайте значения влажности почв (задание 1.3).

Зная сумму осадков и сумму температур, рассчитать коэффициент увлажнения для своего варианта (задание 1.5). Сохранить результаты как скрипт.

Задание 3.2 С помощью функции *seq* (начало, конец, шаг) создать последовательность для своего варианта и записать ее в переменную *p*. Показать содержимое переменной *p*. Создать новую переменную *p1*, в которой последовательность *p* повторяется несколько раз. С помощью команды *c* составить любую простую фразу по-английски из 3-4 слов. Прокомментировать все результаты. Сохранить результаты как скрипт.

Задание 3.3 А. Выполнить операции: $x=c(0,1, 2, 7, 10, 14, 3, 89)$; $x[6]$; $x[3:8]$; $x[c(4, 7)]$; $x[-3]$; $x[-(2:5)]$; $x[x>3]$; $x[x>3 \& x<70]$. Б. Выполнить аналогичные операции для своего варианта. Прокомментировать результаты и сохранить скрипт.

Задание 3.4 Построить следующие графики. Гистограмма для заданной переменной. Гистограмма обилия (выраженное в суммарном количестве растений, встретившихся на угодье (задание 1.4) для видов сорняков, где $ri > 0,03$. Подписать столбцы гистограммы по оси x по-русски. Построить график распределения средних температур в июне по дням по данным, полученным для г. Воронеж в задании 1.5. Подписать название графика, окрасить точки и линии цветными согласно варианту. . Прокомментировать результаты и сохранить скрипт. Сохранить графики.

Раздел 4. Графики и базы данных в среде R

Этап 1. Фреймы

Что такое фрейм данных. Специальные переменные: бесконечность, отсутствие значений и другие. Создание фрейма данных. Обращение к переменной в фрейме данных.

Этап 2. Формат csv и ввод данных

Загрузка данных из внешних источников (диска). Параметр *header*. Параметр *dec*. Как указать рабочую директорию. Команды *summary*, *dim*, *names*. Данные, загруженные по умолчанию. Просмотр загруженных в R данных. Выбор строк, столбцов, ячеек. Логические условия и задание порядка.

Этап 3. Параметры графиков

Задание вида и размера маркера, обозначающего точку на графике. Задание границ осей. Типы линий и задание. Ширина линий. Задание нескольких графиков на странице. Параметры текста: вид шрифта, размер букв и их цвет. Круговая диаграмма. Работа с пропущенными данными.

Задания для самостоятельной работы: Каждое задание выполняется студентом для своего варианта, для каждого задания выполняется отчет.

Задание 4.1 Повторить все команды из лекции из раздела «Фреймы». Письменно прокомментировать их. Для данных своего варианта из задания 1.3 сохранить данные в формате csv в файле. Загрузить содержимое файла в новую переменную. Проверить размерность новой переменной, распечатать ее краткое содержимое (*summary*) и имена столбцов. Рассчитать влажность почвы поэтапно, используя промежуточные вектора. Записать полученный результат в переменную *Moisture*. Сравнить с результатами, полученными в программе Excel.

Задание 4.2 Повторить все команды из раздела «Формат csv и ввод данных». Письменно прокомментировать их. Придумать свой вариант для каждой команды, взяв любую переменную из *data()*, например, *iris* или *prk*. Повторить все команды из раздела «Параметры графиков». Письменно прокомментировать их.

Задание 4.3 Зайти на сайт библиотеки ВАСХНИЛ. Найти в базу Agros. Выбрать проблематику, например, «почва». Скачать количество публикаций по данной проблематике. Построить гистограмму по количества статей по данной тематике в разные временные промежутки. Подписать оси.

Раздел 5. Знакомство с пакетами в среде R, ориентированными на научные расчеты в почвоведении и экологии

Этап 1. Доступные источники информации: книги и пакеты прикладных программ (библиотеки) – общие сведения

Книги, рекомендуемые студентам. R как высокоуровневый объектно ориентированный язык программирования. Типы объектов. Объекты, предназначенные для хранения данных. Программы. Отличия между пакетами и библиотеками. Получение полной информации о пакете. Подключение библиотек и пакетов. Базовые пакеты.

Этап 2. Пакет vegan

Инсталляция пакета. Общие сведения о пакете. Возможности пакета. Примеры использования. Встроенные данные.

Этап 3. Пакет AQP

Алгоритмы для количественного почвоведения. Инсталляция пакета. Общие сведения о пакете. Возможности пакета. Примеры использования. Встроенные данные.

Задания для самостоятельной работы: Каждое задание выполняется студентом для своего варианта, для каждого задания выполняется отчет.

Задание 5.1 Напишите в скрипте напротив имени каждой библиотеки, для чего каждая предназначена: *pastecs; vcd; lattice; gplots; ggplot2; fitdistrplus; reshape; extrafont; psych; knitr; multcomp; HH; GPArotation*. Добавить в список 3-4 библиотеки, связанных с курсовой работой.

Задание 5.2 Познакомиться с файлом *diversity-vegan.pdf*. Знать ответы на следующие вопросы: Когда была написана последняя версия пакета *vegan* и его описания? Чему посвящен данный пакет? Кто его автор? В целом, что содержит описание пакета?

По данным задания 1.4 создать фрейм, где содержится частота встречаемости видов сорняков на поле. Рассчитайте индексы Шеннона и Пиелу.

Задание 5.3 Пакет AQP- Алгоритмы для количественного почвоведения. Знать ответы на следующие вопросы: Когда была написана последняя версия пакета и его описания? Чему посвящен данный пакет? Кто его автор? В целом, что содержит описание пакета?

7. Фонд оценочных средств (ФОС) для оценивания результатов обучения по практике:

7.1. Типовые контрольные задания или иные материалы для проведения текущего контроля успеваемости:

Задания для самостоятельной работы: Каждое задание выполняется студентом для своего варианта, для каждого задания выполняется отчет.

Задание 1.1 Оформление авансового отчета по результатам гипотетической поездки заграницу с пересчетом сумм в рублях и в иностранной валюте.

Задание 1.2 Вычисление потенциальной урожайности (т/га) по формуле, в зависимости от запаса продуктивной влаги в 1 м слое, ожидаемого количества осадков за вегетационный период, непроизводительного расхода почвенной влаги и урожая на единицу израсходованной воды.

Задание 1.3 Расчет весовой влажности для точек с одного поля, расположенных случайным образом. Рассчитать характеристики, используемые для свертки информации: среднее, минимум, максимум, а также число наблюдений. Провести условное форматирование значений влажности для величин, больших влажности завядания (15%). Подсчитать число значений >15, воспользовавшись функцией СЧЕТЕСЛИ(массив, условие).

Задание 1.4 Расчет индексов видового разнообразия сорной растительности для одного из участков угодья. Рассчитать следующие характеристики для варианта: индекс Шеннона, индекс Пиелу.

Задание 2.1 Распределение гранулометрического состава по профилю серых лесных почв. Построить и отформатировать гистограммы распределения фракций гранулометрического состава серых лесных почв по разным глубинам (два типа гистограмм: простые и составные), построить диаграммы типа «точечная», построить круговые диаграммы для отдельных горизонтов. Всего должно быть 10 графиков или более, нужно использовать разные макеты графиков и разные цвета.

Задание 2.2 Расчет Распределение по профилю серых лесных почв подвижных форм железа и микроэлементов в виде кривых распределения по глубинам отбора образцов и в виде гистограмм по горизонтам /подгоризонтам.

Задание 2.3 Расчет сверток и построение графиков для свойств пахотного горизонта на разных элементах рельефа. Рассчитать минимум, максимум и медиану (центральное значение). Построить график, вставьте линии диапазона (Макет – Линии - Коридор диапазона) и отформатировать его.

Задание 3.1 Для 2 строк из авансового отчета (задание 1.1) рассчитать сумму в рублях, исходя из курса валюты и затрат. Прокомментировать результаты. Для двух образцов почв из своего варианта рассчитайте значения влажности почв (задание 1.3). Зная сумму осадков и сумму температур, рассчитать коэффициент увлажнения для своего варианта (задание 1.5). Сохранить результаты как скрипт.

Задание 3.2 С помощью функции *seq* (начало, конец, шаг) создать последовательность для своего варианта и записать ее в переменную *r*. Показать содержимое переменной *r*. Создать новую переменную *r1*, в которой последовательность *r* повторяется несколько раз. С помощью команды *c* составить любую простую фразу по-английски из 3-4 слов. Прокомментировать все результаты. Сохранить результаты как скрипт.

Задание 3.3 А. Выполнить операции: $> x=c(0,1, 2, 7, 10, 14, 3, 89); > x[6]; > x[3:8] ; > x[c(4, 7)] ; > x[-3]; > x[-(2:5)] ; > x[x>3] ; > x[x>3 \& x<70]$. Б. Выполнить аналогичные операции для своего варианта. Прокомментировать результаты и сохранить скрипт.

Задание 3.4 Построить следующие графики. Гистограмма для заданной переменной. Гистограмма обилия (выраженное в суммарном количестве растений, встретившихся на угодье (задание 1.4) для видов сорняков, где $ri > 0,03$. Подписать столбцы гистограммы по оси x по-русски. Построить график распределения средних температур в июне по дням по данным, полученным для г. Воронеж в задании 1.5. Подписать название графика, окрасить точки и линии цветными согласно варианту. . Прокомментировать результаты и сохранить скрипт. Сохранить графики.

Задание 4.1 Повторить все команды из лекции из раздела «Фреймы». Письменно прокомментировать их. Для данных своего варианта из задания 1.3 сохранить данные в формате csv в файле. Загрузить содержимое файла в новую переменную. Проверить размерность новой переменной, распечатать ее краткое содержимое (*summary*) и имена столбцов. Рассчитать влажность почвы поэтапно, используя промежуточные вектора. Записать полученный результат в переменную *Moisture*. Сравнить с результатами, полученными в программе Excel.

Задание 4.2 Повторить все команды из раздела «Формат csv и ввод данных». Письменно прокомментировать их. Придумать свой вариант для каждой команды, взяв любую переменную из *data()*, например, *iris* или *npk*. Повторить все команды из раздела «Параметры графиков». Письменно прокомментировать их.

Задание 4.3 Зайти на сайт библиотеки ВАСХНИЛ. Найти в базу Agros. Выбрать проблематику, например, «почва». Скачать количество публикаций по данной проблематике. Построить гистограмму по количества статей по данной тематике в разные временные промежутки. Подписать оси.

Задание 5.1 Напишите в скрипте напротив имени каждой библиотеки, для чего каждая предназначена: *pastecs*; *vcd*; *lattice*; *gplots*; *ggplot2*; *fitdistrplus*; *reshape*; *extrafont*;

psych; knitr; multcomp; HH; GPArotation. Добавить в список 3-4 библиотеки, связанных с курсовой работой.

Задание 5.2 Познакомиться с файлом *diversity-vegan.pdf*. Знать ответы на следующие вопросы: Когда была написана последняя версия пакета *vegan* и его описания? Чему посвящен данный пакет? Кто его автор? В целом, что содержит описание пакета?

По данным задания 1.4 создать фрейм, где содержится частота встречаемости видов сорняков на поле. Рассчитайте индексы Шеннона и Пиелу.

Задание 5.3 Пакет AQP- Алгоритмы для количественного почвоведения. Знать ответы на следующие вопросы: Когда была написана последняя версия пакета и его описания? Чему посвящен данный пакет? Кто его автор? В целом, что содержит описание пакета?

8. Ресурсное обеспечение:

- Перечень основной и дополнительной литературы:

№ п/п	Автор	Название книги/статьи	Отв. редактор	Место издания	Издательство	Год издания	Название журнала	Том (выпуск) журнала	Номер журнала
1	Шипунов А. Б., Балдин Е. М., Волкова П. А., Коробейников А. И., Назарова С. А., Петров С. В., Суфиянов В. Г.	Наглядная статистика. Используем R			http://ashipunov.info/shipunov/school/books/rbook.pdf	2014			
2	Мастицкий С.Э., Шитиков В.К.	Статистический анализ и визуализация данных с помощью R			http://r-analytics.blogspot.ru/2014/12/r.html#.V16L2TzZQSG	2014			
3	Кабаков Р. И.	R в действии. «Анализ и визуализация данных на языке R»		Москва	ДМК Пресс	2014			
Дополнительная									
4	Афифи А., Эйзен С.	Статистический анализ. Подход с применением ЭВМ		Москва	Мир	1982			
5	Самсонова В.П.	Пространственная изменчивость почвенных свойств. На примере дерново-подзолистых почв		Москва	ЛКИ	2008			
	Пифо Х.-П.	Статистика для бакалавров по специальностям АБ, АН и ВПР в Университете Хоэнхайм		Москва	Изд. ВНИИА	2011			
Интернет-ресурсы									
	https://sites.google.com/site/soilsstatistics/								

	https://office.com						
	http://thermo.karelia.ru/weather						
	http://cran.r-project.org/bin/windows/base/						
	http://www.rstudio.com/ide/download/						
	http://ecologymodeling.github.io/install						
	https://www.datacamp.com						
	http://r-analytics.blogspot.ru						
	http://ropensci.org/						

- Перечень лицензионного программного обеспечения (при необходимости)
- Перечень профессиональных баз данных и информационных справочных систем
- Перечень ресурсов информационно-телекоммуникационной сети «Интернет» (при необходимости)
- Описание материально-технического обеспечения:

А. Помещения

- Лекционная потоковая аудитория, оборудованная оргтехникой (проектор, компьютер, выход в Интернет):
- машинные классы (2 класса с 20 ЭВМ) для проведения статистических расчетов

Б. Оборудование:

Для лабораторных и лекционных аудиторий: необходимая оргтехника, ЭВМ,

В. Иные материалы

9. Язык преподавания: русский

Рабочая программа дисциплины (модуля) разработана в соответствии с самостоятельно установленным МГУ образовательным стандартом (ОС МГУ) для реализуемых основных профессиональных образовательных программ высшего образования по направлению подготовки / специальности «Почвоведение» программы бакалавриата, магистратуры, реализуемых последовательно по схеме интегрированной подготовки в редакции приказа МГУ от 30 декабря 2016 г.