

**Федеральное государственное бюджетное образовательное учреждение
высшего образования
Московский государственный университет имени М.В. Ломоносова
Географический факультет**

УТВЕРЖДАЮ
Декан географического факультета,
член-корр. РАН Добролюбов С.А.

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
Решение прикладных геоморфологических задач с использованием основ
программирования**

Уровень высшего образования:
магистратура

Направление подготовки:
05.04.02 «География»

Направленность (профиль) ОПОП:
«Геоморфология и палеогеография»

Форма обучения:
очная

Рабочая программа рассмотрена и одобрена
Учебно-методической комиссией географического факультета
(протокол № 11 , дата 3 декабря 2021 г.)

Москва 2021

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 «География».

ОС МГУ утверждены решением Ученого совета МГУ имени М.В. Ломоносова (приказ по МГУ № 1383 от 30 декабря 2020 года).

Год (годы) приема на обучение: 2021

1. Место дисциплины (модуля) в структуре ОПОП — относится к вариативной части ОПОП, является дисциплиной по выбору.
2. Входные требования для освоения дисциплины (модуля), предварительные условия: знания, полученные в бакалавриате по математике и основам программирования, основам изучения ГИС, морфометрического анализа.
3. Планируемые результаты обучения по дисциплине (модулю), соотнесенные с требуемыми компетенциями выпускников

Компетенции выпускников (коды)	Индикаторы (показатели) достижения компетенций	Планируемые результаты обучения по дисциплине (модулю), сопряженные с компетенциями
<p>МПК-1 (<i>формируется частично</i>) Способен применять методологию, теоретические концепции геоморфологии и палеогеографии; сопоставлять данные, полученные с помощью системы различных методов изучения морфологии, генезиса, динамики и эволюции рельефа для решения теоретических и прикладных задач.</p>	<p>МПК-1.2 сопоставляет данные, полученные с помощью системы различных методов изучения морфологии, генезиса, динамики и эволюции рельефа для решения теоретических и прикладных задач</p>	<p>Знать: основные принципы устройства языка R, типы данных и наиболее употребляемые для их обобщения базовые функции R; соотношение возможностей для пространственного анализа, даваемых традиционными ГИС-системами с графическим интерфейсом пользователя, и языками программирования на примере R; способы визуализации данных в R; границы применимости основных методов математической статистики, анализа и моделирования при работе с эмпирической геоморфологической и географической информацией;</p> <p>Уметь: создавать собственные («нетривиальные») функции для обработки и анализа данных; разрабатывать сложные многоэтапные расчетные алгоритмы; пользоваться основными методами анализа данных, как «каноническими» в геоморфологии, так и передовыми; визуализировать результаты расчета с помощью базовых функций R, а также пакетов “ggplot2”, “raster”, “rasterVis” и др.; подготавливать корректные наборы геоморфологических данных для различных видов математической обработки и анализа инструментами языка R;</p> <p>Владеть: основами использования языка R, основными пакетами и функциями, применяемыми в геоморфологических исследованиях; методами построения математических моделей, дающих формальное описание геоморфологических объектов и процессов или позволяющих выполнять реконструкции их генезиса и прогноз их принадлежности к определенной группе объектов или прогноз развития, соответственно;</p>

				И И				
Язык R. Особенности использования, синтаксиса.	8	2	4			6	2	2
Наборы геоморфологических данных	6	2	2			4	2	2
<i>Текущая аттестация 1</i>	4		2			2	2	2
Подготовка данных и простейшие статистические операции над ними	5	2				2	3	3
Пространственный анализ	5	2				2	3	3
<i>Текущая аттестация 2</i>	5		2			2	3	3
Регрессионный анализ	5	2				2	3	3
<i>Текущая аттестация 3</i>	5		2			2	3	3
Кластерный анализ	5	2				2	3	3
<i>Текущая аттестация 4</i>	5		2			2	3	3
Классификация с обучением	7	2	2			4	3	3
<i>Текущая аттестация 5</i>	5		2			2	3	3
Подведение итогов	4	4				4		
Промежуточная аттестация зачет	3	<i>Устный зачет</i>						3
Итого	72	18	18			36		36

Содержание лекций, семинаров

Содержание лекций

1. Общие сведения о среде программирования и анализа данных R. Установка Rи клиента RStudio. Основные пакеты, используемые для анализа данных в геоморфологии. Типы данных языка R: векторы, матрицы, списки, таблицы. Упорядоченные серии наблюдений (временные ряды и др.). Типы переменных и их примеры из геоморфологии. Организация вычислений: функции, циклы. Построение графиков.

2. Наборы данных. Чтение и запись различных форматов геоданных в R. Описательные статистики, анализ характера распределения данных. Применимость параметрических и непараметрических методов статистики. Способы предобработки данных (стандартизация, логарифмирование и др., решение проблемы пропущенных значений, снижение размерности). Простейшие статистические тесты для проверки гипотез о равенстве средних и однородности дисперсий в двух распределениях. Использование критериев Хи-Квадрат и Колмогорова-Смирнова. Введение в однофакторный дисперсионный анализ (ANOVA). Критерий Краскела-Уоллиса.
3. Пространственный анализ. Работа с растровыми и векторными данными в R. Извлечение пространственных различий количественных и качественных переменных: зональные статистики, таблицы сопряженности.
4. Регрессионный анализ. Линейная регрессия и корреляция. Обобщенные регрессионные модели. Пробит- и логит-регрессия.
5. Методы кластеризации данных. Иерархическая кластеризация. K-means и fuzzyK-means методы. Случайные деревья и случайный лес. Самоорганизующиеся карты Кохонена. Выбор оптимального набора предикторов для проведения кластеризации.
6. Подготовка классификаций с обучением. Линейный дискриминантный анализ (ЛДА). Использование ЛДА в геоморфологии (примеры из опубликованных работ). Распознавание образов. Машины опорных векторов. Конволюционные нейронные сети.
7. Возможности применения статистического моделирования и методов анализа данных в геоморфологических исследованиях. Описательные и прогностические модели, их отличия, преимущества и слабые стороны. «Классическая» статистика (анализ таблиц наблюдений с информацией о рельефе и геологических особенностях территории) и пространственная статистика (закономерности в пространственных распределениях геоморфологических процессов и объектов). Резюме по лекционной части курса.

Содержание семинаров

1. Освоение принципов языка Rи работы в IDE-клиенте RStudio. Написание первой программы. Рассмотрение основных типов данных и структур языка. Обсуждение отличий разных видов переменных (с примерами из геоморфологии).
2. Ключевые используемые функции. Написание собственных «нетривиальных» функций.
3. Создание набора данных. Освоение основных принципов их компоновки. Знакомство с типами структур данных «vector», «matrix», «data.frame», «z-orderedobservations».
4. Освоение основных пакетов и функций для пространственного анализа. Картометрический анализ векторных геоморфологических и других тематических карт, растровых данных.
5. Построение модели линейной регрессии, прогноз отклика на ее основе (на конкретных геоморфологических данных). Построение прогнозной модели принадлежности объекта к одному из двух классов на основе логистической регрессии.
6. Кластеризация объектов (на примере речных бассейнов). Оценка информативности предикторов и анализ различий групп / кластеров.
7. Освоение некоторых методов классификации с обучением. Построение ансамбля моделей на одном и том же наборе данных, сравнение эффективности разных методов. Обсуждение причин отличий их эффективности.

7. Фонд оценочных средств для оценивания результатов обучения по дисциплине (модулю):

Текущая аттестация 1. Домашнее задание, заключающееся в поиске и адаптации набора данных в опубликованных источниках, или же подготовке собственного набора данных по результатам студенческих курсовых работ и т.п.

Текущая аттестация 2. Домашнее задание, включающее в себя отчет о результатах пространственного анализа, а также защита этого отчета в форме устного доклада с презентацией.

Текущая аттестация 3. Домашнее задание, включающее в себя построение регрессионной модели и оценку ее надежности, а также краткий (не более 1-2 страниц) отчет по результатам этой работы, устное сообщение.

Текущая аттестация 4. Домашнее задание, включающее в себя построение модели кластеризации и оценку ее надежности, а также краткий (не более 1-2 страниц) отчет по результатам этой работы, устное сообщение.

Текущая аттестация 5. Домашнее задание, включающее в себя построение модели классификации и оценку ее надежности, а также краткий (не более 1-2 страниц) отчет по результатам этой работы, устное сообщение.

Промежуточная аттестация: зачет в устной форме

Примерный перечень вопросов

1. Основные понятия языка R, типы данных.
2. Основные виды переменных, примеры из геоморфологии.
3. Ключевые пакеты, используемые при анализе географической и геоморфологической информации в R. Их назначение.
4. Применимость параметрических и непараметрических методов. Способы предобработки данных, необходимые для этого пакеты и функции.
5. Простейшие статистические тесты, их назначение. Примеры возможного применения.
6. Тест t-Стьюдента и однофакторный дисперсионный анализ. Критерий Краскела-Уоллиса.
7. Работа с растровыми и векторными геоданными в R, возможности и ограничения (в сравнении с традиционными ГИС-системами).
8. Виды регрессионного анализа, пример применения линейной регрессии в геоморфологии.
9. Логистическая регрессия, пример ее применения в геоморфологии.
10. Кластеризация данных. Отличия иерархических и неиерархических ее видов. Ограничения метода k-means и способы их обхода. Пример применения в геоморфологии.
11. Случайные деревья и случайный лес (отличия и примеры возможного использования).
12. Самоорганизующиеся карты Кохонена.

13. Классификация с обучением. Линейный и квадратический виды дискриминантного анализа. Примеры из геоморфологии.
14. Классические подходы к распознаванию образов в геоморфологии. Передовые подходы.
15. Линии и поверхности тренда, алгебраические и тригонометрические тренды.
16. Классический и сингулярный спектральный анализ цифровых моделей рельефа. Методы свертки.
17. Численный анализ временных рядов
18. Классическая статистика и геостатистика в геоморфологии.
19. Интерпретационные (описательные) и прогностические модели.
20. Способы выбора оптимального набора предикторов и оценки качества моделей.
21. Написание собственных функций, их тестирование, способы сокращения расчетного времени.

Шкала и критерии оценивания

Текущая аттестация №1. Домашнее задание.

Подготовка тестового набора данных. Конкретный характер и содержание данных – по предварительному согласованию с преподавателем (таблицы наблюдений, пространственные данные и пр.).

Текущая аттестация считается пройденной в случае выполнения следующих требований:

- в электронном виде (формат *.xlsx, *.csv или *.txt) представлен набор данных, который уже используется или потенциально может использоваться автором в решении своих научных или прикладных задач (потребуется устное обоснование);
- группировка данных в наборе позволяет приступить непосредственно к анализу (т.е. нет структурных ошибок);
- сделано устное сообщение о созданном наборе, типе переменных, дано их описание (например, для номинативных переменных – уровни переменной, для интервальных – описательные статистики);
- суммарная оценка должна составлять не менее 66%.

Рубрикатор вклада различных частей задания в суммарную максимальную оценку:

Часть задания	Вклад в суммарную максимальную оценку
Файл набора данных	34%
Отсутствие структурных ошибок в наборе	33%
Устное сообщение	33%
ИТОГО	100%

Текущая аттестация №2. Домашнее задание.

Проведение пространственного анализа. Территория и тип анализируемых данных – на выбор преподавателя.

Текущая аттестация считается пройденной в случае выполнения следующих требований:

- в электронном виде (формат *.docx) представлен текст отчета о проведении пространственного анализа, содержащий разделы: Введение, Исходные данные и проверяемая гипотеза, Результаты анализа, Заключение. Объем отчета – 3-5 стр.
- представленный отчет содержит как численные (желательно – табличные), так и графические результаты, оформленные с помощью R.
- сделан устный доклад по результатам проведенного исследования, сопровождаемый презентацией, подготовленной в электронном виде.
- суммарная оценка за все компоненты задания (текст, картографические материалы, доклад, презентация) должна составлять не менее 50% от максимальной, а оценки за каждый отдельный компонент работы (текстовая часть, картографические материалы, доклад, презентация) должны быть более 0%

Рубрикатор вклада различных частей задания в суммарную максимальную оценку:

Часть задания	Вклад в суммарную максимальную оценку
Текстовая часть отчета	50%
Отражение результатов в табличной и графической формах	25%
Презентация	12.5%
Доклад	12.5%
ИТОГО	100%

Текущая аттестация №3. Домашнее задание.

Разработка модели какого-либо геолого-геоморфологического объекта или процесса (по согласованию с преподавателем) на основе регрессионного анализа.

Текущая аттестация считается пройденной в случае выполнения следующих требований:

- создана регрессионная модель (результат демонстрируется преподавателю в виде скрипта *.R). За использование логистической регрессии + 10% к максимальной оценке.
- сформулировано уравнение регрессии (возможно, без взаимодействия предикторов, за модель с взаимодействием + 5% к максимальной оценке). Оценена значимость предикторов.
- сделано устное сообщение по результатам проведенного исследования, сопровождаемое презентацией (обязательные пункты сообщения — входные данные и их генеральные черты, уравнение регрессии, R^2 , значимость предикторов, прогноз на основе модели).

- суммарная оценка за все компоненты задания должна составлять не менее 50% от максимальной, а оценки за каждый отдельный компонент работы должны быть более 0%.

Рубрикатор вклада различных частей задания в суммарную максимальную оценку:

Часть задания	Вклад в суммарную максимальную оценку
Скрипт, содержащий модель	40%
Уравнение	25%
Устное сообщение	35%
ИТОГО	100%

Текущая аттестация №4. Домашнее задание.

Разработка модели подразделения на группы каких-либо геолого-геоморфологических объектов (по согласованию с преподавателем) на основе кластерного анализа.

Текущая аттестация считается пройденной в случае выполнения следующих требований:

- созданы две модели кластеризации одного и того же набора данных – иерархическими и не иерархическими методами (результат демонстрируется преподавателю в виде скрипта *.R).
- результаты кластеризации визуализированы соответствующими каждому методу способами, построены графики.
- сделано устное сообщение по результатам проведенного исследования, сопровождаемое презентацией (обязательные пункты сообщения — входные данные и их генеральные черты, сравнение результатов кластеризации разными методами, прогноз на основе модели).
- суммарная оценка за все компоненты задания должна составлять не менее 50% от максимальной, а оценки за каждый отдельный компонент работы должны быть более 0%.

Рубрикатор вклада различных частей задания в суммарную максимальную оценку:

Часть задания	Вклад в суммарную максимальную оценку
Скрипт, содержащий обе модели	35%
Графики	20%
Скрипт для осуществления прогноза	20%
Устное сообщение	25%
ИТОГО	100%

Текущая аттестация №5. Домашнее задание.

Модель дифференциации каких-либо геолого-геоморфологических объектов (по согласованию с преподавателем) на основе классификации с обучением.

Текущая аттестация считается пройденной в случае выполнения следующих требований:

- создана модель классификации с использованием одного из методов, по выбору преподавателя (результат демонстрируется преподавателю в виде скрипта *.R).
- результаты классификации визуализированы, построены графики.
- сделано устное сообщение по результатам проведенного исследования, сопровождаемое презентацией (обязательные пункты сообщения — входные данные и их генеральные черты, оценка надежности классификации по тестовой выборке и методом кросс-валидации, анализ информативности предикторов, прогноз на основе модели).
- суммарная оценка за все компоненты задания должна составлять не менее 50% от максимальной, а оценки за каждый отдельный компонент работы должны быть более 0%.

Рубрикатор вклада различных частей задания в суммарную максимальную оценку:

Часть задания	Вклад в суммарную максимальную оценку
Скрипт, содержащий модель	35%
Графики	20%
Скрипт для осуществления прогноза	20%
Устное сообщение	25%
ИТОГО	100%

Успешное выполнение заданий текущего контроля успеваемости

В случае успешного выполнения всех заданий текущего контроля успеваемости и получения за каждое из них оценки не ниже 80% выставляется «зачет».

В случае успешного выполнения всех заданий текущего контроля успеваемости и получения за каждое из них оценки от 50% до 80% проводится устный опрос.

В случае, если какое-либо задание текущего контроля успеваемости не было представлено в полном объеме, или оценка за его выполнение не превысила 50%, выставляется «незачет».

Оценка РО исоответствующие виды оценочных средств	Незачет	Зачет
Знания (виды оценочных средств: устный опрос, домашняя работа)	Фрагментарные знания или отсутствие знаний	Сформированные систематические знания или общие, но не структурированные знания
Умения (виды оценочных средств: устный опрос, домашняя работа)	В целом успешное, но не систематическое умение или отсутствие умений	Успешное и систематическое умение или в целом успешное, но содержащее отдельные пробелы умение (допускает неточности не принципиального характера)
Навыки (владения, опыт деятельности) (виды оценочных средств: устный опрос, домашняя работа)	Наличие отдельных навыков или отсутствие навыков	Сформированные навыки (владения), применяемые при решении задач или, в целом, сформированные навыки (владения), но используемые не в активной форме

8. Ресурсное обеспечение:

- Перечень основной и дополнительной учебной литературы.

Основная литература

1. Кабаков Р.И. R в действии. Анализ и визуализация данных на языке R. М.: ДМК-Пресс, 2014. 588 с.
2. Рыжов П.А., Гудков В.М. Применение математической статистики при разведке недр. М.: Недра, 1966. 236 с.
3. Симонов Ю.Г. Морфометрический анализ рельефа. М.-Смоленск: Изд-во СГУ, 1998. 272 с.
4. Симонов Ю.Г. Объяснительная морфометрия рельефа. Монография. М.: Геос, 1999. 251 с.
5. Шипунов А.Б. и др. Наглядная статистика. Используем R! М.: ДМК-Пресс, 2017. 298 с.

Дополнительная литература:

1. Берлянт А.М. Картографический метод исследования. М.: Изд-во Моск. ун-та, 1978. — 257 с.
2. Коган Р.И., Белов Ю.П., Родионов Д.А. Статистические ранговые критерии в геологии. — М.: Недра, 1983. — 136 с.
3. Рельеф Земли и математика / Под ред. А. С. Девдариани. М.: Мысль, 1967. 117 с.
4. Bivand R.S., Pebesma E., Gómez-Rubio V. Applied Spatial Data Analysis with R. New York: Springer-Verlag, 2013. 405 p.
5. Brunson C., Comber L. An Introduction to R for Spatial Analysis and Mapping SAGE Publications Ltd, 2015. 360 p.
6. Cole J.P., King C.A.M. Quantitative Geography. Techniques and Theories in Geography. — Glasgow: John Wiley & Sons Ltd, 1968. — 692 p.
7. Doornkamp J.C., King C.A.M. Numerical analysis in geomorphology. An introduction. — London: Edward Arnold (Publishers) Ltd., 1971. — 372 p.
8. Dorman M. Learning R for Geospatial Analysis. Packt Publishing, 2014. 364 p.

9. Florinsky I. Digital Terrain Analysis in Soil Science and Geology. Academic Press, 2016. 506 p.
10. King C.A.M. Techniques in Geomorphology. — L.: Edward Arnold, 1966. — 342 p.
11. Spatial analysis in geomorphology / R.J.Chorley. — London: Methuen&Co Ltd., 1972. — 393 p.
12. Taylor P.J. Quantitative Methods in Geography. — Boston: Houghton Mifflin, 1983. — 386 p.

- Перечень лицензионного программного обеспечения

1. R-3.5.2 - язык программирования R (набор основных библиотек).
2. RStudio-1.1.463 - графический интерфейс пользователя для языка программирования R.

- Перечень профессиональных баз данных и информационных справочных систем

The Comprehensive R Archive Network. Адрес доступа: <https://cran.r-project.org/>

- Перечень ресурсов информационно-телекоммуникационной сети «Интернет» (при необходимости)

1. - Шитиков В.К., Мастицкий С.Э. (2017) Классификация, регрессия и другие алгоритмы DataMining с использованием R. 351 с. — Электронная книга, адрес доступа: <https://github.com/ranalytics/data-mining>
2. RStudio – Open source and enterprise-ready professional software for R. Адрес доступа: <https://www.rstudio.com/>
3. Stack Overflow - Where Developers Learn, Share, & Build Careers. Адрес доступа: <https://stackoverflow.com/>

- Описание материально-технической базы

Компьютерный класс, жидкокристаллический экран для демонстрации материалов.

9. Язык преподавания: русский

10. Преподаватель: С.В.Харченко

11. Разработчики программы: Харченко Сергей Владимирович, старший научный сотрудник кафедры геоморфологии и палеогеографии, кандидат географических наук