

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

УТВЕРЖДАЮ
Проректор по учебной работе и
образовательным инновациям
О.Г. Прохоренко
«05» декабря 2022



Регистрационный № УД – 11345/уч.

**МАТЕМАТИЧЕСКИЕ МЕТОДЫ ОБРАБОТКИ И АНАЛИЗА
ГЕОДАНЫХ**

**Учебная программа учреждения высшего образования
по учебной дисциплине для специальности:**

1-56 80 01 Землеустройство, кадастры, геодезия и геоматика

Профилизация: Геоматика

2022 г.

Учебная программа составлена на основе образовательного стандарта высшего образования второй ступени (магистратура) ОСВО 1-56 80 01-2019 специальности 1-56 80 01 Землеустройство, кадастры, геодезия и геоматика, учебного плана БГУ № I56-027/уч. от 11.04.2019 г.

СОСТАВИТЕЛЬ:

А.А. Карпиченко, доцент кафедры почвоведения и геоинформационных систем Белорусского государственного университета, кандидат географических наук, доцент.

РЕЦЕНЗЕНТ:

А.В. Таранчук, заведующий кафедрой географии и методики преподавания географии УО «Белорусский государственный педагогический университет имени Максима Танка», кандидат географических наук, доцент.

РЕКОМЕНДОВАНА К УТВЕРЖДЕНИЮ:

Кафедрой почвоведения и геоинформационных систем факультета географии и геоинформатики БГУ
(протокол № 6 от 24.11.2022 г.);

Научно-методическим Советом БГУ
(протокол № 2 от 29.11.2022 г.)

Заведующий кафедрой _____

А.Н. Червань

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

Актуальность изучения учебной дисциплины «Математические методы обработки и анализа геоданных», которая изучается студентами на второй ступени высшего образования по специальности 1-56 80 01 Землеустройство, кадастры, геодезия и геоматика, обусловлена необходимостью понимания алгоритмов анализа географических данных, широким применением одномерного и многомерного анализов в современных научных и прикладных исследованиях, реализацией многих статистических методов в прикладных ГИС-программах.

Цели и задачи учебной дисциплины

Цель учебной дисциплины «Математические методы обработки и анализа геоданных» – формирование и закрепление знаний, умений и навыков в области применения математической статистики и теории вероятностей при обработке и научном анализе географических данных.

Задачи учебной дисциплины:

- изучение прикладных основ теории вероятности, математической статистики и геостатистики;
- формирование навыков статистической обработки и анализа географических данных с использованием специализированного программного обеспечения;
- приобретение умения решения научно-исследовательских и прикладных задач в области геоматики с использованием среды статистических вычислений R.

Место учебной дисциплины в системе подготовки специалиста с высшим образованием.

Учебная дисциплина «Математические методы обработки и анализа геоданных» относится к **модулю** «Геоинформационное картографирование» компонента учреждения высшего образования.

Учебная программа составлена с учётом межпредметных **связей** с учебными дисциплинами «Методология и современные проблемы геоматики» и «Программное ГИС-обеспечение».

Требования к компетенциям

Освоение учебной дисциплины «Математические методы обработки и анализа геоданных» обеспечивает формирование следующей **специализированной компетенции:**

СК–7. Быть способным анализировать современное состояние и перспективы развития рынка экологических услуг, оценивать экономическую эффективность экологических услуг в научно-инновационной деятельности.

В результате освоения учебной дисциплины магистрант должен:

знать:

- понятийно-терминологический аппарат теории вероятностей;
- основные законы распределения случайных величин в природных и техногенных системах;
- методы одно- и многомерного анализа данных;
- ведущие математико-статистические методы, применяемые при анализе временных рядов;
- теорию пространственной статистики.

уметь:

- правильно использовать терминологию в области теории вероятности и математической статистики;
- производить первоначальный анализ экспериментальных и статистических данных, выдвигать и проверять статистические гипотезы (с учетом характера распределения выборок);
- научно-обоснованно выбирать соответствующий статистический метод обработки и анализа географических данных;
- применять методы регрессионного анализа, многомерного статистического анализа и классификации геоданных;
- осуществлять анализ временных рядов;
- выполнять анализ пространственно-распределенной информации методами геостатистики.

владеть:

- математическими методами решения типовых исследовательских задач при анализе геоданных;
- программными средствами статистической обработки экспериментальной географической информации.

Структура учебной дисциплины

Дисциплина изучается во II семестре. Всего на изучение учебной дисциплины «Математические методы обработки и анализа геоданных» отведено 100 часов, в том числе 48 аудиторных часов, из них: лекции – 24 часа (в том числе 12 ч/ДО), лабораторные занятия – 24 часа (в том числе 10 ч/ДО).

Трудоемкость учебной дисциплины составляет 3 зачетных единицы.
Форма текущей аттестации – экзамен.

СОДЕРЖАНИЕ УЧЕБНОЙ ДИСЦИПЛИНЫ

Раздел 1. Основы анализа геоданных

Тема 1.1. Особенности применения математических методов в геоматике. Географические данные, объект и предмет изучения, типовые решаемые задачи. Базовый понятийный аппарат анализа данных. Работа с данными. Количественные и качественные данные. Основные шкалы измерения. Основные задачи анализа и интерпретации данных. Этапы анализа данных. Программное обеспечение, применяемое для анализа данных..

Тема 1.2. Основные положения теории вероятностей. Случайное событие, виды случайных событий. Операции над случайными событиями. Определение вероятности случайного события. Теоремы сложения и умножения вероятностей событий. Условная вероятность. Формула полной вероятности и формула Байеса. Повторение испытаний. Схема Бернулли. Теоремы Пуассона и Муавра-Лапласа. Случайные величины и их виды: дискретные и непрерывные случайные величины. Закон распределения вероятностей случайной величины. Функция и плотность распределения вероятностей случайной величины, ее свойства. Числовые характеристики случайных величин, их свойства. Виды теоретических распределений. Закон больших чисел и центральная предельная теорема теории вероятностей.

Тема 1.3. Одномерный статистический анализ. Логические основания проверки статистических гипотез. Вычисление показателей описательной статистики. Показатели распределения выборки. Параметрические критерии проверки гипотез. Непараметрические критерии проверки гипотез. Метод Краскала. Медианный тест. Дисперсионный анализ.

Раздел 2. Методы анализа многомерных данных

Тема 2.1. Множественный корреляционно-регрессионный анализ. Представления о многомерном пространстве и размерности. Многомерные распределения случайных событий. Составление регрессионной модели и параметрический регрессионный анализ. Многомерная регрессионная модель. Нелинейная регрессия. Методы структурного моделирования.

Тема 2.2. Методы снижения размерности данных. Метод главных компонент. Математическая модель метода. Геометрическая интерпретация метода главных компонент. Сущность методов факторного анализа и их классификация. Общий вид линейной модели и основные задачи факторного анализа. Общий алгоритм и этапы работ при факторном анализе. Геометрическое представление наблюдаемых объектов в пространстве элементарных признаков и латентных факторов. Многомерное непараметрическое шкалирование.

Тема 2.3. Методы классификации многомерных наблюдений. Классификация без обучения. Кластерный анализ. Расстояние между объектами и мера близости. Меры близости между кластерами. Иерархические процедуры кластеризации. Метод К-средних. Классификация в пространстве главных компонент и общих факторов. Критерии качества классификации. Зависимость выбора метода классификации от цели исследования. Математическое описание метода дискриминантного анализа. Обучающие выборки. Линейный дискриминантный анализ при известных параметрах многомерного нормального закона распределения. Вероятность ошибочной классификации с помощью дискриминантной функции. Оценка качества дискриминантной функции и информативности отдельных признаков. Пошаговый дискриминантный анализ. Деревья решений. Анализ данных с использованием деревьев решений. Критерии и функции качества разбиения узлов дерева. Индекс Джини. Энтропия. Алгоритмы построения деревьев решений. ID3/C4.5, CART, Random Forest. Практические реализации методов построения деревьев классификации с использованием программных средств.

Раздел 3. Временные ряды наблюдений

Тема 3.1. Анализ временных рядов. Компоненты временного ряда. Выделение тренда. Сезонные и циклические колебания. Сглаживание временных рядов. Динамические модели с распределенными лагами. Стационарные временные ряды. Тестирование стационарности. Прогнозирование временных рядов. Адаптивные и мультипликативные методы прогнозирования. Экспоненциальное сглаживание. Авторегрессионные модели. Модели скользящего среднего. Прогнозирование с использованием моделей временных рядов.

Раздел 4. Пространственная статистика и геостатистика

Тема 4.1. Основы пространственной статистики. Пространственная статистика. Пространственная переменная. Случайный процесс в пространстве и его моменты. Первый момент – математическое ожидание. Моменты второго порядка. Дисперсия. Ковариация. Стационарность и эргодичность.

Тема 4.2. Геостатистический анализ. Основные задачи вариограммного анализа. Вариограмма. Основные параметры вариограммы: порог, ранг, наггет-эффект и их интерпретация. Модели вариограмм. Вариограмма по направлению. Анизотропия. Анализ и интерпретация графиков вариограммы по направлениям. Моделирование вариограммы. Выявление и обработка глобальных трендов в данных. Кригинг и базовые модели геостатистики. Обычный кригинг. Схема интерполяции методом кригинга. Точечный и блочный кригинг. Кросс-

валидация. Неопределенность оценки. Понятие о других видах кригинга. Теория кокригинга. Элементы вероятностного картирования. Индикаторный кригинг. Обработка результатов индикаторного кригинга. Описание пространственной вариабельности и неопределенности. Стохастическое моделирование.

Тема 4.3. Пространственная автокорреляция. Пространственная автокорреляция. Анализ географического соседства. Соседство по смежности. Соседство по метрике. Матрица пространственных весов. Показатели пространственной автокорреляции. Глобальные и локальные показатели. Join Count Statistic. Индекс Морана I (Moran's I). Диаграмма рассеяния Морана, Джири С (Geary's C), Getis-Ord's G. Интерпретация результатов. Проверка статистической значимости. Локальные показателей пространственной ассоциации» (LISA) для оценки кластеризации.

Тема 4.4. Географически взвешенная регрессия. Пространственная гетерогенность. Локальная регрессия. Географически взвешенная регрессия. Весовые функции. Фиксированная весовая функция. Адаптивная весовая функция. Полоса пропускания. Интерпретация результатов географически взвешенной регрессии.

УЧЕБНО-МЕТОДИЧЕСКАЯ КАРТА УЧЕБНОЙ ДИСЦИПЛИНЫ

Дневная форма получения образования с применением электронных средств обучения (ДО)

Номер раздела, темы	Название раздела, темы	Количество аудиторных часов					Количество часов УСР	Форма контроля знаний
		Лекции	Практические занятия	Лабораторные занятия	Семинарские занятия	Иное		
1	Основы анализа геоданных	4 2(ДО)		2 4 (ДО)				
1.1	Особенности применения математических методов в геоматике	2						
1.2	Основные положения теории вероятностей	2(ДО)		2 2 (ДО)				Опрос по теме Проверка расчетно-графических работ (РГР)
1.3	Одномерный статистический анализ	2		2 (ДО)				Проверка РГР
2	Методы анализа многомерных данных	6 2(ДО)		6 2 (ДО)				
2.1	Множественный корреляционно-регрессионный анализ	2		2				Устный опрос
2.2	Методы снижения размерности данных	2 2(ДО)		2				Проверка РГР

2.3	Методы классификации многомерных наблюдений	2		2 2 (ДО)				Тестирование Проверка РГР
3	Временные ряды наблюдений	2(ДО)		2				
3.1	Анализ временных рядов	2(ДО)		2				Проверка РГР
4	Пространственная статистика и геостатистика	2 6(ДО)		4 4 (ДО)				
4.1	Основы пространственной статистики	2(ДО)		2(ДО)				Проверка РГР
4.2	Геостатистический анализ	2		2				Тестирование Проверка РГР
4.3	Пространственная автокорреляция	2(ДО)		2				Устный опрос Проверка РГР
4.4	Географически взвешенная регрессия	2(ДО)		2(ДО)				Проверка РГР
	ИТОГО	24		24				

ИНФОРМАЦИОННО-МЕТОДИЧЕСКАЯ ЧАСТЬ

Перечень основной литературы

1. Самсонов Т.Е. Визуализация и анализ географических данных на языке R / Т.Е. Самсонов. – М.: Географический факультет МГУ, 2017. DOI: 10.5281/zenodo.901911.
2. Чертко, Н.К. Математические методы в географии: учеб.-метод. пособие / Н.К. Чертко, А.А. Карпиченко. – Минск: БГУ, 2009. – 199 с.
3. Comber, L. Geographical Data Science and Spatial Data Analysis. An Introduction in R / L. Comber, C. Brunson. – SAGE Publications Ltd., 2020. – 360 p.

Перечень дополнительной литературы

4. Анализ данных в экологии сообществ и ландшафтов; пер. с англ. под ред. А.Н. Гельфана, Н.М. Новиковой, М.Б. Шадринной. – М.: РАСХН, 1999. – 306 с.
5. Боровиков, В.П. Популярное введение в современный анализ данных в системе STATISTICA. Учебное пособие для вузов / В.П. Боровиков. – Москва: Горячая Линия–Телеком, 2018. – 288 с.
6. Геоestatистика: теория и практика / В. В. Демьянов, Е. А. Савельева; под ред. Р. В. Арутюняна; Ин-т проблем безопасного развития атомной энергетики РАН. — М.: Наука, 2010. — 327 с.
7. Мастицкий, С.Э. Статистический анализ и визуализация данных с помощью R / С.Э. Мастицкий, В.К. Шитиков // R-анализ и визуализация данных [Электронный ресурс]. – 2014. – Режим доступа: <http://r-analytics.blogspot.com>
8. Наглядная статистика. Используем R! / А.Б. Шипунов [и др.]. – М.: ДМК Пресс, 2014. – 298 с.
9. Пузаченко, Ю.Г. Математические методы в географических и экологических исследованиях / Ю.Г. Пузаченко. – М., 2004. – 410 с.
10. Садовникова, Н.А. Анализ временных рядов и прогнозирование. Вып. 3: Учебно-методический комплекс / Н.А. Садовникова, Р.А. Шмойлова. – М.: Изд. центр ЕАОИ, 2009. – 264 с.
11. Симчера, В.М. Методы многомерного анализа статистических данных: учеб. пособие / В.М. Симчера. – М.: Финансы и статистика, 2008. – 400 с.
12. Филандышева, Л.Б. Статистические методы в географии: учебно-методическое пособие / Л.Б. Филандышева, Е.С. Сапьян. – Томск, 2015. – 164 с.
13. Чертко, Н.К. Математические методы в географии [Электронный ресурс]: учебное пособие / Н.К. Чертко, А.А. Карпиченко. –

Электрон. текстовые данные. – Саратов: Ай Пи Ар Медиа, 2019. – 193 с. – Режим доступа: <http://www.iprbookshop.ru/84871.html>.

14. Шитиков, В.К. Классификация, регрессия и другие алгоритмы Data Mining с использованием R / В.К. Шитиков, С.Э. Мастицкий [Электронный ресурс]. – 2017. – Режим доступа: <https://github.com/ranalytics/data-mining>.

15. Acevedo, M.F. Data Analysis and Statistics for Geography, Environmental Science, and Engineering / M. F. Acevedo. – CRC Press, 2019. – 560 p.

16. Dadson, S.J. Statistical Analysis of Geographical Data: An Introduction / Simon J. Dadson. – Wiley-Blackwell, 2017. – 264 p.

17. Mario, F. Triola. Elementary statistics / F. Triola Mario. – Pearson, 2017. – 770 p.

18. Qian, S.S. Environmental and Ecological Statistics with R / S.S. Qian. – Chapman & Hall, 2022. – 560 p.

19. Rogerson, Peter A. Statistical Methods for Geography. A Student's Guide / Peter A. Rogerson. – SAGE Publications Ltd., 2019. – 432 p.

ПЕРЕЧЕНЬ РЕКОМЕНДУЕМЫХ СРЕДСТВ ДИАГНОСТИКИ И МЕТОДИКА ФОРМИРОВАНИЯ ИТОГОВОЙ ОТМЕТКИ

Устный опрос: оценка за ответы на лекциях (опрос) и лабораторных занятиях включает полноту ответа, наличие аргументов, примеров из практики и т.д.

Оценка степени усвоения теоретического материала проверяется путем тестирования. При оценке расчетных заданий принимается во внимание правильность проведенных расчетов, грамотное использование нормативно-справочной информации. При оценивании лабораторных работ принимается во внимание корректность полученных результатов, владение соответствующим теоретическим материалом, ответы на контрольные вопросы, способность учащегося теоретически обосновать и детально пояснить полученные результаты и практическую реализацию задания.

Итоговая оценка знаний студента производится по 10-ти балльной шкале и формируется на основе документов:

1. Правила проведения аттестации студентов, курсантов, слушателей при освоении содержания образовательных программ высшего образования (Постановление Министерства образования Республики Беларусь от 29 мая 2012 г. № 53);

2. Положение о рейтинговой системе оценки знаний обучающихся по учебной дисциплине в Белорусском государственном университете (Приказ ректора БГУ № 189 – ОД от 31.03.2020);

3. Критерии оценки результатов учебной деятельности обучающихся в учреждениях высшего образования по десятибалльной шкале (Письмо Министерства образования Республики Беларусь от 28.05.2013 г. № 09-10/53-ПО).

Формой текущей аттестации по дисциплине «Математические методы обработки и анализа геоданных» учебным планом предусмотрен экзамен.

При формировании итоговой отметки используется рейтинговая система оценки знаний студента, дающая возможность проследить и оценить динамику процесса достижения целей обучения. Рейтинговая система предусматривает использование весовых коэффициентов для текущего контроля знаний и текущей аттестации студентов по дисциплине.

Примерные весовые коэффициенты, определяющие вклад текущего контроля знаний и текущей аттестации в итоговую отметку:

Формирование отметки за текущую успеваемость:

– выполнение расчетно-графических работ (среднеарифметическая величина отметок за работы) – 50 %;

- тестирование по темам (среднеарифметическая величина отметок за все тесты) – 40 %

– опрос – 10 %.

Итоговая отметка по дисциплине рассчитывается на основе отметки текущей успеваемости – 40 % и экзаменационной отметки – 60 %.

Описание инновационных подходов и методов к преподаванию учебной дисциплины

Образовательный процесс по учебной дисциплине «Математические методы обработки и анализа геоданных» построен с использованием **практико-ориентированного** подхода, который предполагает:

- освоение содержание образования через решения практических задач;
- приобретение навыков эффективного выполнения разных видов профессиональной деятельности, включая обработку и анализа геоданных при решении научно-исследовательских и практических задач;
- ориентацию на генерирование идей, реализацию групповых проектов, развитие предпринимательской культуры;
- использование процедур, способов оценивания, фиксирующих сформированность профессиональных компетенций.

Методические рекомендации по организации самостоятельной работы обучающихся

Самостоятельная работа студентов по учебной дисциплине «Математические методы обработки и анализа геоданных» предполагает углубленное изучение основной и дополнительной литературы, выполнение исследовательских проектов. Для организации самостоятельной работы студентов по учебной дисциплине предполагается использование современных информационных технологий, размещение в сетевом доступе комплекса учебных и учебно-методических материалов (учебно-программные материалы, учебные издания для теоретического изучения дисциплины, методические указания к лабораторным занятиям, материалы текущего контроля и текущей аттестации (вопросы для подготовки к зачету, задания, тесты), список рекомендуемой литературы, информационных ресурсов и др.

Примерный перечень вопросов к экзамену

1. Основные понятия теории вероятностей. Пространство элементарных событий. Алгебра событий, основные законы событий.
2. Методы задания вероятностей. Определение вероятностей. Геометрический метод задания вероятностей.
3. Формула полной вероятности и формула Байеса.
4. Понятие случайной величины. Виды случайных величин. Дискретные случайные величины (ДСВ). Закон распределения ДСВ.
5. Непрерывные случайные величины. Плотность распределения. Равномерное, логарифмическое, нормальное распределение.
6. Одномерный статистический анализ. Описательные статистики.

7. Графические методы анализа данных.
8. Проверка статистических гипотез. Нулевая и альтернативная гипотезы. Ошибки I и II рода. Статистический критерий проверки нулевой гипотезы. Критическая область.
9. Сравнение выборок. Параметрические и непараметрические методы сравнения двух выборок.
10. Корреляционный анализ. Реализация в R.
11. Линейный регрессионный анализ. Реализация в R.
12. Множественный регрессионный анализ (уравнение множественной регрессии, коэффициенты регрессии, коэффициенты детерминации и пошаговые методы, условия получения приемлемых результатов анализа).
13. Классификация методов многомерного статистического анализа.
14. Сущность методов факторного анализа. Общий алгоритм выполнения. Метод главных компонент.
15. Общая характеристика методов кластерного анализа. Расстояние между объектами и мера близости. Меры близости между кластерами. Иерархические кластер-процедуры.
16. Неиерархический кластерный анализ, метод k-средних.
17. Геостатистический анализ. Основные понятия геостатистики. Этапы геостатистического анализа.
18. Пространственная автокорреляция. Анализ географического соседства. Матрица пространственных весов.
19. Индексы пространственной автокорреляции (Морана I (Moran's I), Джири С (Geary's C), Getis-Ord's G), их интерпретация, проверка статистической значимости).
20. Анализ временных рядов наблюдений.

**ПРОТОКОЛ СОГЛАСОВАНИЯ УЧЕБНОЙ ПРОГРАММЫ
ПО УЧЕБНОЙ ДИСЦИПЛИНЕ «МАТЕМАТИЧЕСКИЕ МЕТОДЫ
ОБРАБОТКИ И АНАЛИЗА ГЕОДАННЫХ» С ДРУГИМИ
ДИСЦИПЛИНАМИ СПЕЦИАЛЬНОСТИ**

Название учебной дисциплины, с которой требуется согласование	Название кафедры	Предложения об изменениях в содержании учебной программы УВО по учебной дисциплине	Решение, принятое кафедрой, разработавшей учебную программу (с указанием даты и номера протокола)
Методология и современные проблемы геоматики	Кафедра почвоведения и ГИС	Нет предложений об изменениях	Вносить изменения не требуется (протокол № 6 от 24.11.2022 г.)

**ДОПОЛНЕНИЯ И ИЗМЕНЕНИЯ К УЧЕБНОЙ ПРОГРАММЕ
«МАТЕМАТИЧЕСКИЕ МЕТОДЫ ОБРАБОТКИ И АНАЛИЗА
ГЕОДАННЫХ»
на ____/____ учебный год**

№ п/п	Дополнения и изменения	Основание
1		
2		

Учебная программа пересмотрена и одобрена на заседании кафедры почвоведения и геоинформационных систем (протокол № _ от _____ 202_ г.)

Заведующий кафедрой
кандидат сельскохозяйственных наук, доцент _____ А.Н. Червань

УТВЕРЖДАЮ
Декан факультета
кандидат географических наук, доцент _____ Е.Г. Кольмакова