

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ БІОРЕСУРСІВ І
ПРИРОДОКОРИСТУВАННЯ УКРАЇНИ

Кафедра таксації лісу та лісового менеджменту

«ЗАТВЕРДЖУЮ»

Директор ННІ лісового і садово-паркового
господарства

П. Лакида

2023 р.



«СХВАЛЕНО»

на засіданні кафедри таксації лісу та лісового
менеджменту

Протокол № 9 від 5 квітня 2023 р.

Т.в.о. завідувача кафедри

В. Миронюк

«РОЗГЛЯНУТО»

Гарант ОНП 205 Лісове господарство

В. Миронюк

РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Аналітичні методи дослідження лісових екосистем

спеціальність 205 Лісове господарство

освітньо-наукова програма «Лісове господарство»

ННІ лісового і садово-паркового господарства

Розробники: професор, докт. с.-г. наук, проф. Миронюк Віктор Валентинович
(посада, науковий ступінь, вчене звання)

Київ – 2023 р.

1. Опис навчальної дисципліни

Аналітичні методи дослідження лісових екосистем

(назва)

Галузь знань, напрям підготовки, спеціальність, освітній ступінь		
Освітньо-науковий рівень	Доктор філософії	
Спеціальність	205 Лісове господарство	
Освітня програма	205 Лісове господарство	
Характеристика навчальної дисципліни		
Вид	Вибіркова	
Загальна кількість годин	150	
Кількість кредитів ECTS	5,0	
Кількість змістових модулів	2	
Форма контролю	Екзамен	
Показники навчальної дисципліни для денної та заочної форм навчання		
	денна форма навчання	заочна форма навчання
Рік підготовки	2	2
Семестр	3	3
Лекційні заняття	30 год.	8 год.
Практичні, семінарські заняття	30 год.	8 год.
Лабораторні заняття	–	–
Самостійна робота	90 год.	134 год.
Індивідуальні завдання	–	–
Кількість тижневих аудиторних годин для денної форми навчання	4 год.	4 год.

2. Мета та завдання навчальної дисципліни

Мета дисципліни «Аналітичні методи дослідження лісових екосистем» полягає в оволодінні методами обробки емпіричних спостережень на основі сучасних комп'ютерних технологій.

Завданнями вивчення курсу «Аналітичні методи дослідження лісових екосистем» є:

- ознайомлення з основними теоретичними та методичними підходами стосовно збору інформації про лісові екосистеми вибірковими методами;
- оволодіння основними статистичними методами інтерполяції та заміщення пропущених даних спостережень;
- вивчення теорії та практичних прийомів із застосування математичного моделювання;
- засвоєння методики розробки лінійних і нелінійних математичних моделей, у тому числі моделей змішаного ефекту;
- набуття навичок верифікації, інтерпретації та практичного застосування математичних моделей;
- ознайомлення з розрахунком невизначеності математичних моделей, оцінок параметрів рівнянь та зробленого на цій основі прогнозу.

У результаті вивчення навчальної дисципліни студент повинен

знати:

- принципи формування експериментальних баз даних, правила їхнього зберігання, організації доступу й обміну інформацією;
- основні програмні продукти, призначені для статистичного аналізу лісівничої інформації;
- теоретичні положення вибірових досліджень, їхню практичну реалізацію під час дослідження лісових екосистем;
- підходи щодо математичного моделювання з використанням різних типів моделей;
- методику інтерполяції та відновлення пропущених даних;
- теоретичні основи діагностики математичних моделей;

вміти:

- використовувати статистичний пакет R для комплексного аналізу лісівничої інформації;
- здійснювати операції імпорту-експорту різнопланової інформації в середовищі системи R;
- виконувати математичне моделювання за допомогою інструментів системи R;
- реалізовувати лінійний і нелінійний регресійний аналіз;
- використовувати алгоритми системи R для інтерполяції та відновлення пропущених даних, в тому числі геостатистичними методами;
- відобразити результати дослідження лісових екосистем у вигляді високоякісної графічної продукції.

Набуття компетентностей:

інтегральна компетентність (ІК):

- ІК1. Здатність продукувати нові ідеї, розв'язувати комплексні проблеми у лісовому секторі економіки, застосовувати методологію наукової та педагогічної діяльності, а також проводити власне наукове дослідження, результати якого мають наукову новизну, теоретичне та практичне значення.

фахові (спеціальні) компетентності (СК):

- СК02. Здатність застосувати сучасні методології, методи та інструменти емпіричних і теоретичних досліджень лісового господарства, цифрові технології, методи комп'ютерного моделювання, бази даних та інші електронні ресурси, спеціалізоване програмне забезпечення у науковій та освітній діяльності;
- СК07. Здатність виявляти, ставити та вирішувати проблеми дослідницького та/або інноваційного характеру у сфері лісового господарства, оцінювати та забезпечувати якість виконуваних досліджень;

Програмні результати:

- РН03. Використовувати сучасні інструменти та технології пошуку, оброблення та аналізу інформації з проблем лісового господарства, зокрема статистичні методи аналізу даних великого обсягу та/або складної структури, спеціалізовані бази даних та інформаційні системи;
- РН08. Формулювати і перевіряти гіпотези; використовувати для обґрунтування висновків належні докази, зокрема результати теоретичного аналізу, експериментальних досліджень і математичного та/або комп'ютерного моделювання, наявні літературні дані.

3. Програма та структура навчальної дисципліни

ЗМІСТОВНИЙ МОДУЛЬ 1 ОРГАНІЗАЦІЯ АНАЛІТИЧНО-ДОСЛІДНИЦЬКОЇ РОБОТИ

Тема лекційного заняття 1

Організаційно-інформаційне забезпечення досліджень лісових екосистем

Основні типи та структури даних. Класифікація моделей даних. Ієрархічні, мережеві та реляційні моделі даних. Системи управління лісівничими базами даних.

Зберігання інформації у цифровому форматі. Особливості доступу до даних, які знаходяться на мережевих дисках і «хмарних» серверах. Способи спільного доступу й обміну даними.

Файлові формати даних для зберігання дослідницької інформації. Основні формати для зберігання текстової інформації. Електронні таблиці. Способи зберігання растрової інформації.

Використання статистичного пакету R для організації науково-дослідницької роботи.

Тема лекційного заняття 2

Вибірковий метод у дослідженні лісових екосистем

Статистичний метод у дослідженні лісових екосистем. Основні поняття теорії вибіркового дослідження. Генеральна сукупність, обсяг вибірки, первинна одиниця вибірки. Розрахунок інтенсивності та репрезентативності вибірки.

Способи організації вибіркового дослідження. Територіальна основа вибірки. Простий випадковий відбір. Систематична та випадкова вибірка. Стратифікована та кластерна вибірка. Методика багатофазного відбору. Комбінація різних способів відбору в дослідженні лісових екосистем.

Розрахунок обсягу вибірки. Виведення середніх і загальних оцінок для вибірки та генеральної сукупності. Побудова довірчих інтервалів і розрахунок невизначеності оцінок. Обчислення обсягу вибірки під час оцінювання часток.

Реалізація методів статистичного відбору в системі R.

Тема лекційного заняття 3

Планування територіальної схеми вибіркового дослідження

Планування збору інформації на територіальній основі. Способи розташування вибірових одиниць. Регулярні прямокутні та гексагональні сітки. Корегування інтенсивності вибірки, зберігаючи рівномірність розподілу вибірових одиниць на місцевості. Приклади типових схем лісової інвентаризації.

Використання статистичного пакету R для проектування територіальної основи вибірки.

Тема лекційного заняття 4

Інтеграція наземних і дистанційних даних у дослідженні лісових екосистем

Дистанційне зондування Землі як джерело додаткових даних про ліси. Поєднання наземних і дистанційних спостережень. Збір реферативних даних візуальними методами. Поєднання наземних спостережень із спектральними даними перехід від показників на пробі до показників на рівні ландшафту.

Практична реалізація задачі в статистичній системі R.

ЗМІСТОВНИЙ МОДУЛЬ 2 КОМПЛЕКСНИЙ АНАЛІЗ ДОСЛІДНОЇ ІНФОРМАЦІЇ

Тема лекційного заняття 5

Методи апроксимації експериментальних даних і побудови моделей

Вивчення характеру розподілу досліджуваного показника як основа використання статистичного аналізу. Параметричні та непараметричні методи оцінювання. Методи «підгонки» основних законів розподілу випадкової величини.

Лінійна регресія. Парна та множинна регресії. Вибір залежних і незалежних змінних. Підбір параметрів прямої лінії методом найменших квадратів. Дослідження рівняння регресії графічними та статистичними методами. Узагальнені лінійні моделі.

Нелінійна регресія. Методи пошуку параметрів нелінійних математичних моделей. Основні нелінійні моделі росту та їхнє використання в дослідженні лісових екосистем. Аналіз багатофакторних нелінійних регресійних рівнянь. Узагальнені нелінійні моделі.

Аналіз ієрархічних даних. Поняття про фіксований і випадковий ефект. Вибір параметрів рівняння з випадковим ефектом як спосіб відображення мінливості відгука. Використання моделей змішаного ефекту для моделювання показників окремих дерев і лісових насаджень. Вибір факторів для групування дослідного матеріалу.

Логістична регресія. Імовірність та шанс події. Логіт-перетворення для нормування даних. Підбір коефіцієнтів логістичної моделі. Прогнозування на основі логістичних моделей.

Побудова моделей в статистичній системі R.

Тема лекційного заняття 6

Статистичні методи відновлення пропущених даних

Проблема неповних даних і відомі способи її вирішення. Методи обробки пропущених значень. Візуальне дослідження структури пропущених спостережень. Виявлення причин пропуску даних і ефекту від них. Метод найближчих сусідів (k -NN) для відновлення пропущених значень у лісовій інвентаризації.

Методи інтерполяції даних. Поліноміальна апроксимація кількісних показників. Роль лінійної інтерполяції в лісівничих дослідженнях.

Інтерполяція і відновлення пропущених даних у системі R.

Тема лекційного заняття 7

Алгоритми ресемплінгу даних

Методи крос-валідації моделей. Використання методів ресемплінгу для валідації математичних моделей. Leave-one-out та k -Fold крос-валідація. Алгоритм bootstrap для оцінювання невизначеності прогнозу.

Практична реалізація алгоритмів крос-валідації в статистичній системі R.

Тема лекційного заняття 8

Оцінювання адекватності математичних моделей

Статистичні критерії оцінки узгодженості емпіричних і модельних значень залежної змінної. Перевірка адекватності математичної моделі за допомогою F -критерія Фішера. Обґрунтування доцільності ускладнення рівняння додатковими параметрами.

Значущість параметрів регресії. Оцінка помилок параметрів математичної моделі. Використання t -критерія Стюдента для побудови довірчих інтервалів параметрів регресійної моделі.

Дослідження адекватності математичної моделі за залишками. Статистичні та графічні підходи стосовно аналізу розподілу залишків регресійного рівняння. Графічний аналіз математичної моделі.

Статистичні та графічні методи діагностики математичних моделей у системі R.

Назви змістових модулів і тем	Кількість годин												
	тижні	денна форма						Заочна форма					
		усього	у тому числі					усього	у тому числі				
			л	п	лаб	інд	с.р.		л	п	лаб	інд	с.р.
1	2	3	4	5	6	7	8	9	10	11	12	13	
Змістовий модуль 1. Організація аналітично-дослідницької роботи													
Тема 1. Організаційно-інформаційне забезпечення досліджень лісових екосистем	1-2	18	4	4	-	-	10	34	2	2	-	-	30
Тема 2. Вибірковий метод у дослідженні лісових екосистем	3-4	18	4	4	-	-	10	38	2	2	-	-	34
Тема 3. Планування територіальної схеми вибіркового дослідження	5-6	18	4	4	-	-	10	-	-	-	-	-	-
Тема 4. Інтеграція наземних і дистанційних даних у дослідженні лісових екосистем	7-8	18	4	4	-	-	10	-	-	-	-	-	-
Разом за змістовим модулем 1	-	72	16	16	-	-	40	72	4	4	-	-	64
Змістовий модуль 2. Комплексний аналіз дослідної інформації													
Тема 5. Методи апроксимації експериментальних даних і побудови моделей	9-10	23	4	4	-	-	15	34	2	2	-	-	30
Тема 6. Статистичні методи відновлення пропущених даних	11	19	2	2	-	-	15	-	-	-	-	-	-
Тема 7. Алгоритми ресемплінгу даних	12-13	18	4	4	-	-	10	-	-	-	-	-	-
Тема 5. Оцінювання адекватності математичних моделей	14-15	18	4	4	-	-	10	44	2	2	-	-	40
Разом за змістовим модулем 2	-	78	14	14	-	-	50	78	4	4	-	-	70
Усього годин	-	150	30	30	-	-	90	150	8	8	-	-	80

4. Темі практичних занять

№ з/п	Назва теми	Кількість годин
1	Експорт-імпорт баз даних із текстових файлів та електронних таблиць у систему R	4
2	Оцінювання описових статистик на основі піднаборів дослідних даних	4
3	Проектування територіальної основи вибірки	4
4	Посидання польових атрибутивних даних і дистанційних спостережень	4
5	Побудова математичних моделей	4
6	Інтерполяції інформації в базах атрибутивних даних	2
7	Оцінювання точності регресійної моделі методом leave-one-out крос-валідації	4
8	Оцінювання неадекватності математичної моделі методом bootstrap	4
	Разом	30

5. Теми самостійної роботи

№ з/п	Назва теми	Кількість годин
1	Написання користувацьких функцій для системи R	10
2	Створення циклічних операторів в системі R	10
3	Робота векторними типами просторових об'єктів в системі R	10
4	Робота з растровими типами даних в системі R	10
5	Вивчення інструментів системи R для математичного моделювання	15
6	Програмні можливості системи R для відновлення даних	15
7	Написання алгоритму крос-валідації	10
8	Написання алгоритму для bootstrap апроксимації	10
	Разом	90

5. Контрольні запитання

1. Дайте визначення поняття «модель даних».
2. Які існують класифікації моделей даних?
3. Наведіть приклад багатовимірної схеми відображення часових рядів за допомогою гіперкуба?
4. Наведіть приклади ієрархічних моделей даних, які використовуються під час дослідження лісових екосистем.
5. Наведіть приклад реляційних баз даних.
6. Викладіть особливості формування запитів у реляційних базах даних.
7. Опишіть переваги і недоліки зберігання дослідницької інформації у текстових файлах порівняно з електронними таблицями.
8. Основні формати растрової графіки та особливості її використання в дослідженні лісових екосистем.
9. Наведіть основні команди, які дозволяють виконувати імпорт текстової інформації до системи R.
10. Як на мові R виконати запис даних до текстових і растрових форматів?
11. Які існують базові структури системи R.
12. Опишіть основні команди для роботи з векторами, списками і фреймами даних в середовищі R. Як виконати перетворення форматів даних?
13. Опишіть переваги використання вибірових методів під час дослідження лісових екосистем.
14. Охарактеризуйте підходи до формування випадкової і механічної вибірки.
15. Які основні відмінності між кластерною і стратифікованою вибірками?
16. Як розрахувати необхідний обсяг вибірки?
17. Яким чином здійснюється побудова довірчих інтервалів для оцінок середніх значень під час оцінювання часток?
18. Поясніть суть оцінювання на основі вибірки, методу відношень і регресії.
19. Опишіть основні функції системи R для формування вибірки.
20. Поясніть зміст процедури беггінгу. В яких випадках вона використовується?
21. Поясніть відмінність між процедурами інтерполяції та відновлення пропущених даних.
22. Які існують підходи до заміщення пропусків у базах даних?
23. Для яких задач може використовуватися інтерполяція растрової інформації під час дослідження лісових екосистем.
24. Створіть функцію в системі R для виконання лінійної інтерполяції між двома точками.
25. За допомогою яких операторів забезпечуються циклічні розрахунки в системі R?

26. Наведіть класифікацію моделей. Які моделі найпоширеніші в лісовій справі?
27. На які етапи поділяється процес моделювання?
28. Поясніть суть моделей змішаного ефекту. Яким чином повинна бути організована база даних для розробки подібних моделей?
29. Особливості розробки багатофакторних математичних моделей.
30. Опишіть послідовність ускладнення математичної моделі додатковими параметрами. Які критерії при цьому використовуються?
31. Зобразіть схему дисперсійного аналізу.
32. Як обчислюється сума квадратів залишків, неадекватності та чистої помилки?
33. Які функції системи R призначені для виконання лінійного і нелінійного регресійних аналізів?
34. Які статистичні критерії використовуються для оцінки неадекватності математичних моделей?
35. Продемонструйте приклади побудови основних типів графіків у системі R.
36. Які функції використовуються для створення точкових і лінійних графіків у середовищі R?
37. Перерахуйте параметри, які дозволяють налаштувати відображення графічної інформації в R: лінії регресії, легенда, назви осей, колір, маркери, лінії тощо.
38. Програмні можливості R для побудови гістограм.
39. Розміщення кількох графіків на одному аркуші плиткою.
40. Як здійснюється експорт графічної інформації до файлу в системі R?

6. Методи навчання

Під час вивчення дисципліни використовуються словесні, наочні та практичні методи навчання.

7. Форми контролю

Основною формою контролю засвоєння дисципліни є залік. Після завершення вивчення навчального матеріалу в межах кожного змістовного модуля проводиться контроль знань у вигляді тесту. Хід виконання індивідуальних завдання систематично контролюється викладачем під час занять.

8. Розподіл балів, які отримують здобувачі. Оцінювання знань студента відбувається за 100-бальною шкалою і переводиться в національні оцінки згідно з табл. 1 «Положення про екзамен та заліки у НУБіП України» (наказ про уведення в дію від 03.03.2021 р. протокол № 7).

Рейтинг студента, бали	Оцінка національна за результати складання	
	екзаменів	заліків
90-100	Відмінно	Зараховано
74-89	Добре	
60-73	Задовільно	
0-59	Незадовільно	Не зараховано

Для визначення рейтингу студента з навчальної роботи та дисципліни використовуються наступна методика:

Поточний контроль		Рейтинг з навчальної роботи ($R_{НР}$)	Рейтинг з додаткової роботи ($R_{ДР}$)	Рейтинг штрафний ($R_{ШТ}$)	Підсумкова атестація (екзамен) (R_A)	Загальна кількість балів ($R_{Дис}$)
Змістовий модуль 1 (R_{M_1})	Змістовий модуль 2 (R_{M_2})					
0-100	0-100	0-70	0-20	0-5	0-30	0-100

$$R_{НР} = 0,7 \cdot \frac{R_{M_1} \cdot 2,5 + R_{M_2} \cdot 2,5}{5,0} + R_{ДР} - R_{ШТ}$$

$$R_{ДИС} = R_{НР} + R_{АТ}$$

КРИТЕРІЇ ОЦІНЮВАННЯ ЗНАТЬ СТУДЕНТІВ
з дисципліни «Аналітичні методи дослідження лісових екосистем»

Практичні роботи (Самостійна робота)	Кількість балів	Проміжний контроль	Всього
МОДУЛЬ 1	Годин/кредитів ECTS		75/2,5
ПР1. Експорт-імпорт баз даних із текстових файлів та електронних таблиць у систему R <i>СР1. Написання користувацьких функцій для системи R</i>	15		
ПР2. Оцінювання описових статистик на основі піднаборів дослідних даних <i>СР2. Створення циклічних операторів в системі R</i>	15		
ПР3. Проектування територіальної основи вибірки <i>СР3. Робота векторними типами просторових об'єктів в системі R</i>	15		
ПР4. Послання польових атрибутивних даних і дистанційних спостережень <i>СР4. Робота з растровими типами даних в системі R</i>	15		
Разом	60	40	100
МОДУЛЬ 2	Годин/кредитів ECTS		75/2,5
ПР5. Побудова математичних моделей <i>СР5. Вивчення інструментів системи R для математичного моделювання</i>	15		
ПР6. Інтерполяції інформації в базах атрибутивних даних <i>СР6. Програмні можливості системи R для відновлення даних</i>	15		
ПР7. Оцінювання точності регресійної моделі методом leave-one-out крос-валідації <i>СР.7. Написання алгоритму крос-валідації</i>	15		
ПР8. Оцінювання неадекватності математичної моделі методом bootstrap <i>СР.8. Написання алгоритму для bootstrap апроксимації</i>	15		
Разом	55	45	100

9. Методичне забезпечення

1. Моделювання продуктивності лісів : Методичні вказівки для виконання практичних робіт у програмному забезпеченні R / НУБіП України ; розроб. : В. В. Миронюк, С. М. Кашпор, В. А. Свинчук, М. М. Кутя, К., 2015. 30 с.

2. Аналітичні методи дослідження лісових екосистем. ЕНК. ULR:
<https://elearn.nubip.edu.ua/course/view.php?id=4434>

10. Рекомендовані джерела інформації

– основна

1. Robinson A. P., Hamann J. D. Forest Analytics with R. An Introduction. New-York: Springer Science+Business Media, LLC. 2011. 339 p.
2. James G., Witten D., Hastie T., Tibshirani R. An Introduction to Statistical Learning with Application in R. New-York: Springer Science+Business Media, LLC. 2013. 426 p.

– допоміжна

1. Zuur A. F. Mixed effects models and extensions in ecology with R. New York, NY:Springer, 2009. 574 с.
2. Cochran, W.G. Sampling Techniques. 3rd Edition, John Wiley & Sons, New York. 1977. P. 428.
3. Никитин К. Е., Швиденко А. З. Методы и техника обработки лесоводственной информации. М. : Лесн. пром-сть, 1978. 272 с.