

**НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ БІОРЕСУРСІВ І
ПРИРОДОКОРИСТУВАННЯ УКРАЇНИ**

Кафедра таксації лісу та лісового менеджменту

ЗАТВЕРДЖУЮ

Перший проректор Національного університету
біоресурсів і природокористування України,
професор, академік НААН

_____ І. Ібатуллін

« _____ » _____ 20__ р.

РОЗГЛЯНУТО І СХВАЛЕНО

на засіданні науково-технічної ради НДІ
лісівництва та декоративного садівництва
протокол № _____ від _____ 20__ р.
Голова НТР

_____ Р. Васишин

на засіданні кафедри таксації лісу та лісового
менеджменту

протокол № _____ від _____ 20__ р.

Завідувач кафедри

_____ А. Білоус

РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Аналітичні методи дослідження лісових екосистем

спеціальність _____ 205 – «Лісове господарство»

ННІ _____ лісового і садово-паркового господарства

Розробники: _____ доцент, докт. с.-г. наук, доц. Миронюк Віктор Валентинович

(посада, науковий ступінь, вчене звання)

1. Опис навчальної дисципліни

Аналітичні методи дослідження лісових екосистем

(назва)

Галузь знань, напрям підготовки, спеціальність, освітній ступінь		
Галузь знань	20 – Аграрні науки та продовольство	
Спеціальність	205 – «Лісове господарство»	
Освітньо-науковий рівень	«Доктор філософії»	
Характеристика навчальної дисципліни		
Вид	Вибіркова	
Загальна кількість годин	150	
Кількість кредитів ECTS	5,0	
Кількість змістових модулів	2	
Форма контролю	Залік	
Показники навчальної дисципліни для денної та заочної форм навчання		
	денна форма навчання	заочна форма навчання
Рік підготовки	1	1
Семестр	2	2
Лекційні заняття	20 год.	20 год.
Практичні, семінарські заняття	20 год.	20 год.
Лабораторні заняття	–	–
Самостійна робота	110 год.	110 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 і технікою обробки інформації під час дослідження лісових екосистем з урахуванням сучасних тенденцій розвитку комп'ютерних технологій.

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

- ознайомлення з основними теоретичними і методичними підходами стосовно збору інформації про лісові екосистеми вибірковими методами;
- оволодіння основними статистичними методами інтерполяції та заміщення пропущених даних спостережень;
- вивчення теорії та практики математичного моделювання на системній основі, включаючи аналіз взаємозв'язку біометричних ознак;
- 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 технікою розрахунку 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;
- техніку 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;
- відображати результати дослідження лісових екосистем у вигляді високоякісної графічної продукції.

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

Зберігання інформації у цифровому форматі. Особливості обробки даних, які знаходяться на магнітних носіях, 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 вибірових досліджень. Популяція, обсяг вибірки, первинна одиниця вибірки. Розрахунок інтенсивності та репрезентативності вибірки.

Способи організації вибірового дослідження. Територіальна основа вибірки. Простий випадковий відбір. Систематична і випадкова вибірка. Стратифікована і кластерна вибірка. Методика багатофазного відбору. Комбінація різних способів відбору в дослідженні лісових екосистем. Приклади типових схем лісової інвентаризації.

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

Реалізація методів статистичного відбору в системі R.

ЗМІСТОВНИЙ МОДУЛЬ 2 КОМПЛЕКСНИЙ АНАЛІЗ ДОСЛІДНОЇ ІНФОРМАЦІЇ

Тема лекційного заняття 3

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.

Основні поняття і елементи геостатистики. Методи інтерполяції растрових поверхонь. Способи опису просторової мінливості та кореляції даних. Побудова семіваріограм. Інтерполяція поверхонь методом кригінга (*kriging*). Використання геостатистичних методів інтерполяції під час дослідження лісових екосистем.

Інтерполяція і відновлення пропущених даних у системі R.

Тема лекційного заняття 4

Методи апроксимації експериментальних даних і побудови моделей

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

Лінійна регресія. Парна і множинна регресії. Вибір залежних і незалежних змінних. Підбір прямої методом найменших квадратів. Дослідження рівняння регресії графічними і 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 багатofакторних нелінійних регресійних рівнянь. Узагальнені нелінійні моделі.

Аналіз ієрархічних даних. Роль моделей змішаного ефекту під час моделювання продуктивності лісових екосистем. Вибір параметрів рівняння, які мають випадковий ефект, як спосіб відображення мінливості відгука.

Програмні можливості системи R для моделювання в дослідженнях лісових екосистем.

Тема лекційного заняття 5

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												
	денна форма							Заочна форма					
	тижні	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	1–2	16	4	2	–	–	10	16	4	2	–	–	10
Тема 2. Вибірковий метод у дослідженні лісових екосистем	3–4	16	4	2	–	–	10	16	4	2	–	–	10
Тема 3. Статистичні методи відновлення пропущених даних	5–6	28	4	4	–	–	20	28	4	4	–	–	20
Разом за змістовим модулем 1	–	60	12	8	–	–	40	60	12	8	–	–	40
Змістовий модуль 2. Комплексний аналіз дослідної інформації													
Тема 4. Методи апроксимації експериментальних даних і побудови моделей	7–8	30	4	6	–	–	20	30	4	6	–	–	20
Тема 5. Техніка оцінки адекватності математичних моделей	9–10	30	4	6	–	–	20	30	4	6	–	–	20
Разом за змістовим модулем 2	–	60	8	12	–	–	40	60	8	12	–	–	40
Усього годин	–	120	20	20	–	–	80	120	20	20	–	–	80

4. Теми практичних занять

№ з/п	Назва теми	Кількість годин
1	Експорт-імпорт баз даних із текстових файлів та електронних таблиць у систему R	2
2	Базові операції з растровими та векторними типами просторових об'єктів	2
3	Користувачські функції системи R	2
4	Цикли та умовні оператори	2
5	Описові статистики та візуалізація даних	4
6	Побудова регресійних моделей	2
7	Моделі класифікації даних	2
8	Створення інтерактивної візуалізації даних за допомогою пакету {shiny}	4
	Разом	20

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 вибіркою?
 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.
 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 студенти

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 ($R_{М_1}$)	Змістовий модуль 2 ($R_{М_2}$)					
0–100	0–100	0–70	0–20	0–5	0–30	0–100

$$R_{НР} = 0,7 \cdot \frac{R_{М_1} \cdot 2,0 + R_{М_2} \cdot 2,0}{4,0} + R_{ДР} - R_{ШТ}$$

$$R_{дис} = R_{НР} + R_{АТ}.$$

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

Практичні роботи	Кількість балів	Проміжний контроль	Всього
МОДУЛЬ 1	Годин/кредитів ECTS	60/2,0	
1. Експорт-імпорт баз даних із текстових файлів та електронних таблиць у систему R	10		
2. Базові операції з растровими та векторними типами просторових об'єктів	15		
3. Користувацькі функції системи R	15		
4. Цикли та умовні оператори	15		
Разом	55	45	100
МОДУЛЬ 2	Годин/кредитів ECTS	60/2,0	
5. Описові статистики та візуалізація даних	10		
6. Побудова регресійних моделей	15		
7. Моделі класифікації даних	10		
8. Створення інтерактивної візуалізації даних за допомогою пакету {shiny}	20		
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.

10. Рекомендована література

– основна

1. Кабаков Р. И. R в действии. Анализ и визуализация данных на языке R. М.: ДМК Пресс, 2014. 388 с.

2. Шитиков В. К., Мاستицкий С. Э. Классификация, регрессия и другие алгоритмы Data Mining с использованием R. ULR: <https://ranalytics.github.io/data-mining/index.html>.

3. Robinson A. P., Hamann J. D. Forest Analytics with R. An Introduction. New-York: Springer Science+Business Media, LLC. 2011. 339 p.

– допоміжна

1. Громов Ю. Ю., Иванова О. Г., Яковлев А. В., Однолько В. Г. Управление данными. Тамбов: Изд-во ФГБОУ ВПО «ТГТУ», 2015. 192 с.

2. Дрейпер Н., Смит Г. Прикладной регрессионный анализ. М.: Финансы и статистика, 1986. 366 с.

3. Кокрен У. Методы выборочного исследования. М. : Статистика, 1976. 440 с.

4. Никитин К. Е., Швиденко А. З. Методы и техника обработки лесоводственной информации. М. : Лесн. пром-сть, 1978. 272 с.

ПЛАН ЛЕКЦІЙ

Тема лекційного заняття 1

ОРГАНІЗАЦІЙНО-ІНФОРМАЦІЙНЕ ЗАБЕЗПЕЧЕННЯ ДОСЛІДЖЕНЬ ЛІСОВИХ ЕКОСИСТЕМ

1. Основні типи та структури даних.
2. Зберігання, обробка та обмін даними.
3. Файлові формати даних.
4. Статистичний пакет R як універсальний інструмент для обробки даних.

Тема лекційного заняття 2

ВИБІРКОВИЙ МЕТОД У ДОСЛІДЖЕННІ ЛІСОВИХ ЕКОСИСТЕМ

1. Основні поняття теорії вибірових досліджень.
2. Методи випадкового відбору.
3. Розрахунок обсягу вибірки.
4. Аналіз невизначеності вибірових оцінок параметрів.
5. Реалізація методів статистичного відбору в системі R.

Тема лекційного заняття 3

СТАТИСТИЧНІ МЕТОДИ ВІДНОВЛЕННЯ ПРОПУЩЕНИХ ДАНИХ

1. Методи обробки пропущених значень.
2. Техніка інтерполяції.
3. Геостатистичні методи інтерполяції.
4. Інтерполяція та заміщення пропущених даних в системі R.

Тема лекційного заняття 4

МЕТОДИ АПРОКСИМАЦІЇ ЕКСПЕРИМЕНТАЛЬНИХ ДАНИХ І ПОБУДОВИ МОДЕЛЕЙ

1. Апроксимація ймовірнісних розподілів таксаційних показників насаджень.
2. Лінійна регресія.
3. Нелінійна регресія.
4. Побудова моделей змішаного ефекту.
5. Техніка математичного моделювання мовою R.

Тема лекційного заняття 5

ОЦІНЮВАННЯ АДЕКВАТНОСТІ МАТЕМАТИЧНИХ МОДЕЛЕЙ

1. Статистичні критерії для оцінювання неадекватності моделей.
2. Дослідження параметрів регресійного рівняння.
3. Аналіз залишків математичної моделі.
4. Графічні методи в моделюванні лісових екосистем.
5. Діагностика математичних моделей в системі R.