

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Харківський національний університет імені В.Н. Каразіна
Біологічний факультет
Кафедра фізіології та біохімії рослин і мікроорганізмів

Кваліфікаційна робота
здобувача другого (магістерського) рівня вищої освіти
за темою:

ФЕНОЛОГІЯ ТА ПРОСТОРОВЕ РОСПОВСЮДЖЕННЯ PHALLUS IMPUDICUS
І PHALLUS HADRIANI В УКРАЇНІ З ДАНИМИ ВІДКРИТИХ РЕСУРСІВ З
БІОРІЗНОМАНІТТЯ

Спеціальність (спеціалізація) 091 біологія та біохімія
Освітня програма Біологія

Виконавець:



Димитрій Сіварт

Науковий керівник:



Андрій Усиченко

Харків, 2025

РЕФЕРАТ

Робота викладена на 66 сторінках друкованого тексту (гарнітура Times New Roman, шрифт 14 пгт), містить 14 рисунків, та 39 літературних джерела.

Ключові слова: *Phallus impudicus*, *Phallus hadriani*, фенологія, DOY, T30, P30, GAM, SDM, iNaturalist, Україна.

Робота присвячена аналізу фенології та просторового розповсюдження *Phallus impudicus* і *Phallus hadriani* в Україні на основі відкритих спостережень та кліматичних даних. Після фільтрації сформовано 369 валідних записів, до яких додано кліматичні предиктори (T30, P30, ступенедні, висота місцевості).

Метою дослідження було оцінити вплив погодних умов на строки плодоношення та визначити чинники, що формують просторову придатність середовища. Для цього використано кругову статистику, узагальнені адитивні моделі (GAM) та моделювання ареалу (SDM/MaxEnt).

Встановлено, що обидва види плодоносять переважно у другій половині літа (середні DOY ~214–216), а ключовими детермінантами фенології є 30-денна середня температура та кумулятивні опади. Просторові моделі свідчать про нерівномірну реакцію видів на кліматичні умови та вказують на регіони з найвищою потенційною придатністю. Новизна роботи полягає в поєднанні даних громадянської науки з кліматичними предикторами для кількісного аналізу цих видів в Україні.

ЗМІСТ

РЕФЕРАТ	2
ВСТУП	4
РОЗДІЛ 1. Стан досліджень з <i>Phallus</i> spp. та фенології грибів	7
1.1. Загальна фенологія грибів: теорія та підходи	7
1.2. Фенологія гастероміцетів і специфіка <i>Phallus</i> spp	10
1.3. Citizen science у мікології: можливості та обмеження	15
1.4. Методи аналізу фенології і просторового моделювання – огляд	19
1.5. Виклики війни, клімату та рідкісності видів	22
РОЗДІЛ 2. Матеріали та методи досліджень	26
2.1. Джерела даних: запити iNaturalist, “Гриби України” Facebook група	26
2.2. Попередня обробка: фільтрація, дедуплікація, quality criteria	28
2.3. Кліматичні дані: WorldClim/ERA5 – агрегація (30-дн., місяць), degree-days	29
2.4. Обчислення фенологічних змінних: DOY, cumulative degree-days	32
2.5. Статистичні моделі: опис GAM (формула, smoothing choice), LMM/GLMM, circular stats, тести тренду	35
2.6. SDM: MaxEnt – background selection, bias correction, block CV	37
РОЗДІЛ 3. Результати досліджень та обговорення	41
3.1 Опис даних	41
3.2 Сезонність і DOY: графіки, circular analysis, тренди	45
3.3 Результати GAM / LMM: ефекти температури/опадів, partial effect plots	50
3.4 SDM карти придатності, важливість змінних, оцінки (AUC/TSS)	55
ВИСНОВКИ	60
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	62
SUMMARY	66

ВСТУП

Фенологія – наука про календарні біологічні цикли, яка вивчає сезонні події живих організмів. Кліматичні зміни суттєво впливають на фенологічні події багатьох організмів *exempli gratia*, цвітіння рослин та міграції тварин [37]. Однак фенологія грибів залишається порівняно мало вивченою [11;37]. Тим часом гриби є значущим учасниками екосистем – вони беруть участь в кругообігу поживних речовин і утворюють симбіози з рослинами та тваринами. Отже, дослідження їх фенології важливе для прогнозування реакції екосистем на глобальні зміни і розуміння екологічної динаміки.

Для грибів фенологія означає час утворення плодового тіла – видимого органу розмноження, що сигналізує про діяльність міцелію. При цьому відсутність плодових тіл в певний час не означає відсутності гриба – міцелій може приховано рости в ґрунті чи деревині [2]. Тому джерелами фенологічних даних про гриби традиційно є фіксовані спостереження та новітні цифрові бази даних. Історичні колекції вважаються «золотим стандартом» фенологічних досліджень, оскільки вони часто охоплюють довгі часові та широкі просторові шкали [2]. Однак ці дані мають недоліки: нерівномірність збирання, залежність спостерігачів та відсутність інформації про негативні спостереження.

В останні роки для фенології грибів широко використовуються відкриті бази даних біорізноманіття, куди користувач завантажують реєстри зустрічей з грибами. Приклади таких платформ – GBIF (Global Biodiversity Information Facility) та iNaturalist, а також спеціалізовані джерела на кшталт: Mushroom Observer тощо. Ці системи акумулюють сотні тисяч географічно та часово промаркованих записів макроміцетів. Зокрема, національні міко-портали (як данський MycoPortal) містять понад 1,1 млн верифікованих записів грибів, інтегрованих із GBIF [29]. Волонтери й аматори-мікологи фіксують грибні спостереження з фотографіями і прив'язкою до GPS, завантажуючи їх до публічних баз iNaturalist, Mushroom Observer [29]. Хоч ці дані несуть в собі відомі упередження, як от змінна активність спостерігачів і верифікація, вони

дозволяють отримати величезні обсяги інформації майже в реальному часі. Наприклад, кілька досліджень показали, що фенологічні сигнали по типу зрушення дат початку і кінця сезонів плодоношення, узгоджуються при аналізі різних наборів даних, спостережень [2]. Це дає підставу вважати, що масиви “громадських” даних можна використовувати для виявлення довгострокових трендів у фенології грибів [2].

Аналітичні підходи до фенології грибів включають розрахунок стандартних показників сезонності: дату початку, середину та кінець плодоношення протягом року. Наприклад, за методикою з одного з досліджень, датою початку вважають 2,5-й перцентиль річного розподілу дат спостережень, а кінцем – 97,5-й перцентиль [2]. Аналогічно обчислюють середню дату плодоношення. Такий підхід дозволяє порівняти фенологію грибів у різних часових рамках чи регіонах. Існують також альтернативні визначення: наприклад, “екологічний” метод – на основі кумулятивних інтегралів концентрації спор або чисельності плодових тіл [37]. Ще один приклад – нова шкала PhenoFun, яка стандартизує етапи розвитку грибів для узагальненого порівняння фенології [11].

Таким чином, сучасна мікологія отримала можливості фенологічного аналізу грибів: поєднуються традиційні спостереження і великі дані цифрових реєстрів. Теоретично це дозволяє виявляти зв’язки фенології з кліматичними змінними [37] та вивчати відмінності між екологічними групами грибів. Наукова актуальність теми полягає в тому, що розуміння фенології грибів допомагає передбачити реакцію екосистем на глобальні зміни. Наприклад, зміщення сезонів плодоношення може вказувати на адаптацію грибів до потепління чи вологості середовища [2, 37].

За мету роботи було поставлено оцінку фенології та просторове розповсюдження *Phallus impudicus* і *Phallus hadriani* в Україні на основі відкритих геоприв’язаних спостережень і кліматичних даних. Відповідно до мети було поставлено наступні завдання

1. Зібрати геоприв’язані спостереження *Phallus impudicus* і *Phallus hadriani* з відкритих джерел та зібрати пов’язані метадані.

2. Провести первинну обробку даних: фільтрацію записів низької якості, видалення дублікатів і відбір підтверджених (research-grade) спостережень.
3. Обчислити фенологічні змінні (DOY, cumulative degree-days) і підготувати вибірки для аналізу сезонності.
4. Проаналізувати сезонність і її детермінанти з використанням кругової статистики, трендових тестів та GAM/GAMM для кількісної оцінки впливу кліматичних змінних.
5. Побудувати просторові моделі (MaxEnt) для оцінки потенційної придатності середовища, виконати просторову валідацію та підготувати карти розподілу з короткими рекомендаціями щодо інтерпретації й обмежень.

РОЗДІЛ 1

СТАН ДОСЛІДЖЕНЬ З *PHALLUS SPP.* ТА ФЕНОЛОГІЇ ГРИБІВ

1.1 Загальна фенологія грибів: теорія та підходи

Фенологія – це наука про періодичність природних процесів у житті організмів у зв'язку з сезонними змінами клімату [35]. За визначенням найстарішої онлайн енциклопедії: Britannica, фенологія – це дослідження дат повторюваних природних подій у зв'язку із сезонними кліматичними коливаннями, тобто поєднання екологічних та метеорологічних аспектів середовища [35]. Інакше кажучи, фенологія вивчає, коли протягом року настають ті чи інші етапи життєвого циклу організмів у відповідь на зміни температури, освітленості, вологості тощо [11, 35]. В одному дослідженні зазначають наступне: “фенологія є значущою адаптивною ознакою організмів, яка визначає їх біотичні взаємодії з середовищем”, і на підґрунті кліматичних змін ця галузь набуває все більшої важливості [11].

Для грибів фенологічними подіями найчастіше є появи плодових тіл та споруляція. В помірних зонах більшість макроміцетів утворюють плодові тіла насамперед восени, тоді як деякі види пристосувалися плодоносити навесні [18]. Було встановлено, що кліматичні умови одного року корелюють з обсягом та строками плодоношення в наступному сезоні: підґрунтовий міцелій реагує на погодні умови попереднього періоду, що відбивається у датах появи плодового тіла наступного року [18]. Зокрема, спостерігається, що глобальне потепління призвело до зміщення строків весняного плодоношення грибів на більш ранні дати у європейських широтах [18]. Водночас треба враховувати, що плодове тіло гриба існує дуже коротко – зазвичай його можна спостерігати на поверхні ґрунту лише кілька днів на рік [39]. Така епізодичність плодоношення ускладнює збір даних та систематичне спостереження за фенологією грибів [39].

Фенологічні процеси грибів визначаються комплексом факторів. Найважливішими є кліматичні показники: в помірному кліматі визначальним

фактором є температура, тоді як у тропіках вирішальну роль грають опади та вологість [35]. Проте не менше значення мають інші абіотичні умови як от: рН, структура ґрунту, вологість ґрунту, освітленість, натомість відіграють свою роль ще й біотичні компоненти: структура лісових насаджень, наявність рослин-господарів чи симбіонтів [33]. Зокрема, в цьому ж дослідженні, показали, що зміни режимів опадів, температури та водного балансу, а також якість і структура лісового середовища суттєво впливають на продуктивність грибів і час їхнього плодоношення [33]. Таким чином, абіотичні та біотичні умови середовища складають фундамент для фенології макроміцетів.

Через короткочасність появи плодового тіла класичні фенологічні обстеження грибів супроводжуються труднощами. Зазвичай польові дослідження передбачають регламентовані виїзди на фіксацію дат появи плодових тіл разом із аналізом історичних гербарних записів і спостережень. Водночас відсутність довготривалих та систематичних фенологічних даних для багатьох видів призводить до використання непрямих підходів: наприклад, економічна статистика, динаміка ринкових цін на популярні їстівні гриби або національні бази спостережень і колективні наукові платформи залучаються як сурогатні показники продуктивності грибів [39]. Новий напрям полягає в уніфікації фенологічних вимірювань – в одному з вищезгаданих досліджень запропонували фенологічну шкалу PhenoFun для грибів і грибоподібних організмів, аналогічну рослинному BBCH. Ця система кодує послідовність життєвих стадій грибів, що дозволяє стандартизовано фіксувати наявність та тривалість кожної фази їхнього циклу [11].

У сучасних дослідженнях взаємозв'язок фенології з кліматом і екологією виходить на перший план. Багато авторів відзначають зрушення фенологічних патернів, спричинені зміною клімату: наприклад, у Північній Європі було задокументовано раннє плодоношення весняних видів грибів упродовж останніх десятиліть [18]. Такі фенологічні зміни можуть впливати на взаємодію грибів з іншими організмами і на загальний екологічний баланс. Як підкреслюють дослідники, у контексті глобальних кліматичних змін фенологія набуває

важливості соціально-економічного значення [11]. У підсумку, загальне розуміння фенології макроміцетів створює основу для специфічних досліджень сезонності *Phallus impudicus* та *Phallus hadriani*, адже воно дозволяє сформулювати очікування щодо строків їхнього плодоношення та залежностей від навколишніх умов.

Окрім екологічної ролі, фенологічні дослідження грибів важливі й з огляду на їхній біохімічний потенціал. Макроміцети синтезують численні біологічно активні метаболіти, як полісахариди, фенолітотуди, терпенів тощо, це речовини що мають лікарсько-профілактичний ефект. Полісахариди Basidiomycota демонструють імуномодулюючі та протипухлинні властивості [8]. Зокрема, відомо, що такі β -глюкани стимулюють макрофаги та інші імунні клітини й безпосередньо гальмують ріст ракових клітин [8]. Також було показано, що місцеве застосування полісахаридів міцелію *Phallus impudicus* значно прискорює епітелізацію шкірних ран у тваринних моделях [7]. Це свідчить про високий регенеративний потенціал цього виду. Таким чином, фенологічні дослідження грибів мають прикладний вимір: знання про періоди плодоношення дозволяє вчасно збирати плодові тіла з метою отримання цінних біомолекул і водночас зберігати сталий стан популяцій.

Загальна фенологія грибів надає концептуальну базу для вивчення сезонних особливостей плодоношення *Phallus spp.* та їхнього просторового поширення. Розуміння загальних закономірностей часового розподілу плодових тіл, а також екологічних детермінантів фенології дозволить правильно інтерпретувати дані по *Phallus spp.* і сформулювати обґрунтовані гіпотези для подальшого моделювання сезонних та просторових патернів їх появи.

1.2 Фенологія гастероміцетів і специфіка *Phallus spp.*

Гастероміцети демонструють різноманітні пристосування до умов довкілля. Так, “сухоспорові” форми e.g. *Lycoperdon*, *Geastrum* мають щільні шкірки і частіше ростуть у посушливих біотопах – їхні плодові тіла витримують високу температуру і мало випаровують вологу, а спори розсіюються вітром чи дощем [27]. Натомість “слизоспорові” гастероміцети e.g. *Phallus spp.*, *Clathrus*, *Mutinus* сприяють поширенню спор за допомогою комах, приваблюючи їх неприємним запахом глеби. Наприклад, у *Phallus impudicus* розклад органіки стимулює утворення різко пахучих летких сполук, сірчистих, фенольних тощо, що приваблюють мух [15]. Екологічно це означає, що *Phallus spp.* зростають в теплих, вологих умовах: для багатьох видів період плодоношення припадає на теплу пору року та погодні умови після дощу. Зазвичай усі гастероміцети плодоносять із запізненням від настання достатньої вологості та температури, переважно літом–осінню в помірному поясі. Деякі форми, до прикладу, *Geastrum* можуть формувати плодові тіла в пізню осінь і навіть зберігати їх узимку, їх представника можна спостерігати на рисунку 1.1. [13].



Рис. 1. 1. Зображення *Geastrum fimbriatum* [13].

Важливим екологічним чинником є середовище: багато видів *Phallus spp.* та *Mutinus* із *Phallales* є сапротрофами, та проростають на розкладеному деревному субстраті або мульчі, тобто на розкладений органічній підстилці з трісок, кори, листя чи інших решток [15, 28], тоді як низка *Lycoperdon* воліють піщані сухі ґрунти. Морфологічно гастероміцети мають індивідуальні риси розвитку: наприклад, *Phallus spp.* мають дві стадії, власне звичне плодове тіло, а також стадія яйця яка передує стадії розкритого плодового тіла, тобто це означає, що гриб розвиваються із “яйця” з вольвою і розкриваються стрімко, настільки що часто часу треба менше ніж 24 години за сприятливих умов [15], тоді як *Geastrum* поступово розпадаються на “промені” – ніжкоподібні структури, що піднімають споривий мішок. Загалом фенологія цих груп тісно пов’язана з вологоутримуючими властивостями плодових тіл та механізмами розповсюдження спор [27;28].

Спори *Phallus impudicus* і *Phallus hadriani* розсіюються мухами, тож їхні плодові тіла мають яскраво виражений сезонний цикл. У помірному кліматі Європи *P. impudicus* найчастіше плодоносить влітку й на початку осені. Зокрема,

за спостереженнями у Британії стінки вегетативних “яєць” з’являються з червня, а масивне плодоношення триває до жовтня, представника з цього виду можна спостерігати на рисунку 1.2. [25].



Рис. 1. 2. Зображення *Phallus impudicus* [25].

Аналогічно *P. hadriani* – виявляється там же, переважно в піщаних ґрунтах, також з червня до жовтня, його представника можна спостерігати на рисунку 1.3. [24].



Рис. 1. 2. Зображення *Phallus hadriani* [24].

У Центрально-Східній Європі ці строки приблизно збігаються з нашими: *Phallus spp.* ростуть у теплу, вологу пору зазвичай в періоді липень–вересень, і пік плодоношення часто припадає на середину літа. Так, українські натуралісти реєструють *P. impudicus* у листяних та змішаних лісах у другій половині літа, липень–серпень, а *P. hadriani* – переважно влітку на узбережжі Чорного моря. Хоча систематичних фенологічних досліджень саме в Україні бракує, відомо, що у всій Європі ці види плодоносять у схожі терміни за сприятливих температурі й вологості [24;25].

Фенологія *Phallus spp.* є типовою для слизоспорових грибів, але відрізняється від інших гастероміцетів. Наприклад, *Mutinus caninus* плодоносить майже одночасно з *Phallus*: у Британії й Ірландії його вегетаційні тіла з'являються в липні й зникають до жовтня [23]. *Clathrus ruber* також активний влітку – у південній Англії з червня по вересень (у південній Європі цей вид плодоносить майже цілий рік через м'який клімат [10]). Натомість, *Geastrum spp.*

та інші сухоспорові гастероміцети утворюють плодові тіла переважно восени і зимою: їхні перидії плодоносять восени й часто зберігаються на ґрунті всю зиму [13]. Отже, *Phallus spp.* за термінами плодоношення подібні до інших *Phallales*, але випереджають більшість *Geastrum spp.* і трюфелеподібних видів, що розвиваються пізньою осінню [10;13].

Представники роду *Phallus* мають низку екологічних і морфофізіологічних особливостей, які визначають їхню фенологію. По-перше, вони є сапротрофами на багатих гумусом субстратах – наприклад, *P. impudicus* часто асоціюється з дубово-буковими лісами, де багато опалого листя та гниючого дерева [15]. Кореневища з білих ризоморфів прикріплюються до перегною або старої мульчі, що забезпечує необхідну вологість і живлення [15;28]. *P. hadriani*, натомість, спеціалізується на піщаних, заболочених ґрунтах [24]. По-друге, *Phallus spp.* мають спільний механізм розповсюдження спор: їхня глеба є слизистою й надзвичайно ароматною, що залучає до рознесення спор численні мух і жуків [28]. Комахи, пожираючи чи просто контактувавши зі спорами, несуть їх на себе чи у фекаліях у віддалені ділянки, аналогічно до запилення квітів комахами [15;28]. Молекулярний склад глеби, а саме метандіол, фенілоацетальдегід, меркаптани тощо, сприяють цьому зв'язку зі спільниками-ентомофагами. По-третє, швидкий ріст плодового тіла в *Phallus spp.* [15]. дозволяє максимально використати короткі сприятливі періоди вологості. Морфологічно це проявляється в наявності добре вираженої вольви та тканин, що швидко вбирають воду. Наприклад, у *P. impudicus* відомо, що “яйця” можуть проростати у дорослі гриби вже за 4-6 години після інтенсивних опадів [15]. Ці риси – толерантність до вологості, асоціація з мікоризоутворюючими деревами [15] та спеціалізована дисперсія спор – разом формують унікальну фенологію *Phallus spp.* Крім того, хімічний склад плодового тіла також залежить від стадії розвитку: аналіз летких речовин показує, що їхня концентрація зростає під час дозрівання плодової форми і досягає максимуму на останніх стадіях розвитку гриба [30]. Це означає, що фенологічний вік *Phallus impudicus* впливає на доступність його біоактивних компонентів.

Таким чином, фенологія гастероміцетів, зокрема роду *Phallus*, суттєво залежить від їхньої екологічної стратегії та морфології. Зростання плодових тіл у теплий вологий період і взаємодія зі спільниками-інсектами визначають час появи плодових тіл *Phallus spp.* Усвідомлення цих особливостей є важливим для розуміння ролі цих грибів у екосистемах та для використання їх у практиці. Оскільки в плодових тілах *Phallus impudicus* і *Phallus hadriani* виявлено численні біоактивні полісахариди і фенольні сполуки, змінна фенологія може визначати їхню фармакологічну цінність [30]. Відтак дослідження фенології цих видів є ключовим як для екологічних досліджень, так і для практичних застосувань у біотехнології чи медицині.

1.3 Джерела фенологічних даних

У мікології дедалі ширше використовують дані громадянської науки – це маніфестується добровільними спостереженнями грибів, що опубліковані на загальнодоступних платформах. Серед найпопулярніших міжнародних ресурсів – iNaturalist і Mushroom Observer, де користувачі розміщують геомарковані знімки макроміцетів для ідентифікації спільноту [14]. Наприклад, на iNaturalist станом на червень 2023 накопичено понад 8,7 млн спостережень близько 20 тис. видів грибів, проте станом на вересень 2025 накопичено понад 17,421,799 та спостережень близько 26 тис. видів грибів, [14]. Mushroom Observer містить понад 480 тис. спостережень (18,9 тис. таксонів) від учасників з усього світу [14]. Інші платформи – Observation.org, мобільні застосунки та проєкти FunDiS чи FungiMap – також активно долучають аматорів до реєстрації грибів [14;29]. Крім того, ентузіасти фіксують грибні знахідки у соціальних мережах. Так, велика Facebook-група “Mushroom Identification” налічує понад 471,9 тис. учасників, які розмістили понад 1 млн знімків грибів [14]. Хоча такі неофіційні джерела не підпадають під наукові бази даних, вони ілюструють зацікавленість публіки і можуть бути потенційним джерелом інформації. У багатьох країнах також є національні портали (напр. датський Атлас грибів) або проєкти на зразок

“Mushrooms of Keny” які використовують технології спільноти і дані CS для локальних моніторингів [14].

Дані “громадської науки”, далі citizen science (CS) мають низку важливих переваг. По-перше, вони охоплюють широку просторову і часову шкалу. Добровільні спостереження надходять щодня з усього світу, часто значно частіше та з більшою територіальною дисперсією, ніж професійні обліки [29]. У країнах з небагатьма мікологоами CS може заповнювати прогалини в моніторингу: iNaturalist упродовж п’яти років зібрав понад 69 тис. записів про українські гриби, що часто перевищує обсяги традиційних опитувань [29]. По-друге, учасники таких проєктів можуть натрапляти на рідкісні чи нові для науки види. Так, дослідження показали, що зусилля любителів грибів часто призводять до виявлення невідомих або ендемічних макроміцетів [14]. По-третє, CS-дані охоплюють колосальні обсяги: мільйони спостережень дозволяють аналізувати тренди, будувати карти поширення та перевіряти систематику. Наприклад, поєднана база iNaturalist і Mushroom Observer містить понад півмільйона записів лише за Канаду [14]. Такі великі набори даних уже використовували для переперегляду таксонів та доповнення ареалів грибів [14]. По-четверте, CS-ініціативи залучають широку громадськість до науки, підвищуючи екологічну обізнаність і збираючи дані, які професіонали ніколи не охопили б у силі самотужки [29]. Нарешті, накопичені спостереження зі штампом часу та місця дають змогу досліджувати фенологічні цикли і будувати моделі видового поширення (SDM). Наприклад, аналіз різних джерел даних показав, що фенологічні закономірності плодоношення грибів узгоджуються між CS-спостереженнями та традиційними записами [4]. А автоматичне імпортування “research grade” спостережень з iNaturalist до GBIF підвищує повноту карт ареалів грибів, полегшуючи SDM-аналіз.

Разом з перевагами CS-дані мають низку суттєвих обмежень. По-перше, помилки ідентифікації. Оскільки спостерігачі мають різний рівень досвіду, у базах CS можуть опинитися помилково визначені види, особливо дрібні чи мікроскопічні грибки, тобто ті гриби, які важко визначити з фотографії [29].

Часто спостереження публікуються без фізичного екземпляра, що унеможливорює подальшу підтверджену експертизу. По-друге, якість зображень та метаданих. Якщо фото нечітке або без детальних пов'язаних географічних даних, то такі записи втрачають наукову цінність і вимагають ручної фільтрації. По-третє, дані CS нерівномірно розподілені в просторі й часі. Гуртки грибників частіше працюють поблизу населених пунктів і в сезон, коли гриби плодоносять, тому області з рідким населенням або весняні/літні види можуть бути недообліковані. Аналогічно існує таксономічне зміщення: звичні, помітні макроміцети документуються набагато частіше, ніж незрозумілі мікроміцети чи криптичні види, до того ж існує ще ціла низка соціальних факторів яку будуть вносити свою лепту, навіть розвиненість країни, її благополуччя, теж грають роль в кількості спостережень [14, 29]. Все це потребує ретельної фільтрації даних перед аналізом. На щастя, багато платформ передбачають внутрішню систему контролю: наприклад, iNaturalist позначає записи як “research grade” лише за умови достатньої інформації й підтвердження ідентифікації спільнотою [14]. Такі перевірені записи потрапляють у GBIF і використовуються для наукових цілей, що підвищує якість кінцевих висновків [14].

Насиченість відкритих баз даних вже дозволила провести низку наукових досліджень фенології й поширення грибів. Так, в одному дослідженні порівняли фенологічні шаблони плодових тіл грибів за даними різних джерел Швейцарії та Великій Британії. Автори виявили, що попри відмінності у методиках збирання, загальні фенологічні відповіді видів були схожими між базами [4]. Цей приклад показав, що наявні нерепрезентативні спостереження можуть давати узгоджувані результати фенологічних трендів. Інший приклад – датський Атлас грибів – який завдяки понад 1,1 млн експертно перевірених записів макроскопічних грибів дав змогу дослідити споживання деревних субстратів. У США дослідники зібрали й опублікували відкритий набір даних із 400 260 географічно прив'язаних спостережень для 1483 видів грибів у штаті Нью-Джерсі. Цей матеріал уже представлено у GBIF і може слугувати для подальшого картографування ареалів макроміцетів. Хоча прикладів вузькоспеціалізованих

досліджень на конкретних *Phallus spp.* у літературі наразі мало, можна відзначити, що в країнах із розвинутою CS часто використано дані та проєкти, в яких *Phallus spp.* потрапляють до огляду через легкість ідентифікації їхніх плодів. В цілому, наведені приклади демонструють, що зібрані ентузіастами дані успішно застосовують у наукових публікаціях для картографування розподілу грибів і аналізу сезонності їх плодоношення [4].

Оцінюючи надійність громадянських даних, слід зважати на вищеописані обмеження. З одного боку, в умовах браку офіційних обліків мікофлори CS-дані часто є єдиним доступним масштабним джерелом інформації. В Україні, де немає потужних систематичних моніторингів грибів, саме iNaturalist “заповнює прогалину” – надаючи широку базу даних по макроміцетах [29]. Вказано, що iNaturalist уже задокументував понад 2 000 грибних видів з України (понад 69 тис. записів) і служить додатковим інструментом до професійних спостережень [29]. З іншого боку, без спеціальної обробки ці дані складно застосовувати в дослідженнях. Їхню надійність підвищують: 1) використання тільки “research grade” спостережень [14]; 2) перехресна перевірка з гербарними колекціями чи експертними базами; 3) моделювання з урахуванням вибіркового зсувів; 4) залучення професійних мікологів до ревізії великих наборів спостережень. Дослідники відзначають, що при таких обережних підходах результати на основі CS-даних можуть бути досить коректними – зокрема, вони дають узгоджені фенологічні висновки і реалістичні карти розповсюдження [4, 14]. Таким чином, за належної перевірки дані CS можна вважати цінним доповненням до академічних досліджень, особливо коли офіційна інформація недостатня.

Підсумовуючи, громадянська наука має велике значення для вивчення грибів роду *Phallus* в Україні. Умови спостереження за цими макроміцетами – явні плодоношення зі специфічним запахом – сприяють тому, що їх часто фіксують. Натомість офіційних систематичних обліків присвячених саме *Phallus spp.* немає або вони мізерні. Тому CS-дані можуть суттєво доповнити знання про ареали та фенологію *Phallus spp.* у нашій країні [29]. Водночас для академічних висновків слід застосовувати строгі фільтри якості. Якщо ці умови дотримано,

дані громадянської науки стануть потужним ресурсом для майбутніх фенологічних та екологічних досліджень грибів *Phallus* в Україні.

1.4 Методи аналізу фенологічних змін

Узагальнені адитивні моделі (GAM) – це розширення лінійної регресії, що дозволяє нелінійно зв'язувати залежну змінну з предикторами за допомогою плавних функцій GAM і їхні змішані варіанти (GAMM) широко застосовують у фенології грибів, оскільки вони враховують складні просторово-кліматичні закономірності без жорстких лінійних припущень. Наприклад, в огляді одного дослідження. для європейських грибів пояснювалося, що GAMM допускають аналіз без нормальності розподілу і можуть включати рандомні ефекти для різних видів [5]. Гладкі терми в GAM зазвичай реалізуються сплайнами: в описаному дослідженні фенологічні відповіді були змодельовані саме зі сплайновими предикторами просторових координат і висоти, бо “сплайнові предиктори забезпечують ступінь нелінійності, яка краще відповідає географічним закономірностям” [5]. Параметри згладжування у моделей вибираються статистично, як було зроблено при моделюванні присутності грибів за кліматичними змінними [3].

Графічне представлення часткових ефектів – partial effect plots ілюструє вплив одного предиктора на ймовірність появи плодового тіла за умови усереднення інших факторів. Такі графіки показують, як змінюється прогнозована відповідь при зміні одного кліматичного або часово-сезонного показника. У синергізмі з Лінійні змішані моделі (LMM) та GLMM, GAM дозволяють виявляти тонкі криволінійні тренди фенології грибів, які могли б залишатися прихованими при використанні простіших моделей.

LMM та їхні генералізовані аналоги (GLMM) включають одночасно фіксовані та випадкові ефекти. За їх допомогою можна врахувати ієрархічні структури даних – наприклад, різні ділянки та часи обстеження. Так, в іншому дослідженні використали лінійну LMM з вкладеним випадковим ефектом

“блок/ділянка” для оцінки впливу мікроклімату на характеристики плодівниць грибів [19]. Аналогічно, в дослідженні згаданому раніше моделювали середньодобові дати плодоношення з урахуванням видового випадкового ефекту, щоб компенсувати таксономічну неоднорідність реакцій та нерівномірність вибірок [5]. GLMM застосовуються у випадках, коли відповідь не є нормально розподіленою e.g. біноміальні дані “наявність/відсутність”, або Poisson-розподілені лічильні дані плодівниць. Вкладені випадкові ефекти типу: блоки, сезони або регіони дозволяють коректно врахувати просторові чи часові групування й уникнути псевдоспостережень, що критично для аналізу фенології і розподілу грибів [5;19].

Для фенологічних даних час року є циклічною величиною 0–365 днів як коло. Кругова статистика враховує, що наприклад 1-ше і 365-те дні року сусідні на колі. Дані перетворюють на кути – радіани, і обчислюють середній вектор: його напрям – кут, відповідає середній фазі сезону появи плодівниць, а довжина вектора – модулярність, характеризує концентрацію спостережень навколо цього середнього від 0 – рівномірно по року до 1 – всі дати співпадають. Зі статистичних тестів найпоширеніший – тест Рейлі, який перевіряє гіпотезу про однорідний розподіл дат на колі; відхилення від однорідності вказує на наявність яскравої сезонності. Тест Вотсона та інші подібні призначені для порівняння двох або більше кругових вибірок наприклад, сезонності в різних регіонах. Як відзначають дослідники фенології, ігнорування циклічності призводить до помилок у висновках, натомість “кругова статистика є адекватним підходом для обчислення фенологічних параметрів” і коректного порівняння сезонності між системами [34].

При тривалих спостереженнях зміни фенології відображаються у трендах, зсув дат першої появи плодівниць тощо. Для виявлення таких монотонних трендів часто використовують непараметричні тести рангових кореляцій, зокрема тест Манна–Кендалла. Цей тест стійкий до ненормальності даних та ідентифікує статистично значущий зростаючий чи спадний тренд. Наприклад, аналіз історичних даних показав, що ранньо осінні види грибів в

середземноморській зоні нині плодять у середньому на 16–17 діб пізніше, ніж у XIX столітті [36]. Такі контрастні зміщення фенологічних дат чітко виявляються за допомогою трендових тестів. Тобто, тест Манна–Кендалла дозволяє формалізувати зміщення у фенології на фоні кліматичних змін і визначити, чи є вони статистично значущими.

Максимальна ентропія (MaxEnt) – популярний метод екологічного моделювання з наявності виду. Його суть – побудова розподілу ймовірності виду в просторі, який за умови заданих екологічних обмежень має максимальну ентропію, але відповідає відомим точкам присутності. Як описано у, MaxEnt мінімізує відносну ентропію між двома розподілами: емпіричним розподілом середовища у точках знахідок і розподілом середовища на тлі всієї території [38]. Модель використовує відомі місцезнаходження виду та набір екологічних змінних, клімат, ґрунт тощо, а решта простору – як фонові вибірки для оцінки наявності. Важливими налаштуваннями є вибір області прогнозування – extent, і корекція просторових зміщень вибірки (наприклад, введення ваги або “bias file” для компенсування нерівномірності спостережень) [38].

Для валідації SDM часто застосовують просторово-блокову крос-валідацію, розбиваючи регіони на блоки, щоб уникнути псевдокореляцій між тренувальною та тестовою вибірками. У виходах MaxEnt оцінюють важливість змінних: наприклад, “permutation importance” показує, наскільки знижується фіт при випадковому перемішуванні даних однієї змінної. Якість моделі зазвичай вимірюють за ROC-AUC та True Skill Statistic. Високі значення цих метрик (AUC $\rightarrow$ 1, TSS $\rightarrow$ 1) свідчать про гарне передбачення. Наприклад, одна з моделей продемонструвала AUC=0,990 і TSS=0,989 [38], що вказує на високу точність побудованого розподілу виду. Усі ці підходи уже застосовувалися в дослідженнях просторового розподілу макроміцетів і дозволяють

будувати карти потенційного ареалу на основі наявних даних присутності.

Розглянуті методи статистичного аналізу забезпечують всебічну оцінку просторово-сезонної екології макроміцетів. Дослідження показали, що GAMM можуть описувати географічні патерни дат плодоношення грибів та екстраполювати їх на різні регіони Європи [5]. Кругова статистика та трендові тести дозволили виявити послідовні кліматичні зрушення в фенології. середземноморські гриби почали плодити значно пізніше порівняно з попередніми століттями [36]. Методи змішаних моделей корисні для врахування структури даних, а SDM–MaxEnt – для оцінки потенційних ареалів. Таким чином, комбінування GAM, LMM/GLMM, кругової статистики, трендових тестів і SDM дає змогу глибоко вивчити просторово-сезонну динаміку грибних угруповань. Зокрема для *Phallus spp.* ці інструменти допоможуть кількісно охарактеризувати сезонність появи їхніх плодівниць та спрогнозувати поширення у просторі за різних кліматичних умов.

1.5. Виклики війни, клімату та рідкісність видів

Воєнний конфлікт суттєво ускладнив мікологічні дослідження. Обстріли та підриви призвели до масового знищення природних біотопів – лісів, боліт та навіть старих парків, що порушило стабільність екосистем [20]. Пожежі, викликані військовими діями, нищили тисячі гектарів лісу та підривали ґрунтовий шар, позбавляючи гриби місць існування. Окрім прямої руйнації, близько третини території України забруднено мінами й вибухонебезпечними залишками війни [22]. За оцінками уряду, майже 174 тисячі км² є потенційно небезпечними через міни та нерозірвані боєприпаси [22], тому поля, ліси й дюни залишаються недоступними для польових досліджень десятиліттями. Інакше кажучи, відновлення моніторингу грибних ареалів у таких регіонах наразі неможливе без попереднього розмінування.

Крім того, війна руйнує наукову інфраструктуру. Лабораторії, гербарії та інші дослідницькі об'єкти пошкоджено або повністю зруйновано, ускладнивши доступ до зібраних матеріалів та обладнання. Дослідження показали, що воєнні дії призвели до “знищення та забруднення природних середовищ” і спричинили руйнування дослідницьких фондів, що може серйозно ускладнити наукову роботу із відновлення пошкоджених екосистем [20]. Наприклад, вивезення цінних гербарних зібрань з окупованого Херсона показує, що унікальні матеріали опинилися під постійною загрозою знищення; такі операції проходили під обстрілами і лише завдяки ризику для життя дослідників. Таким чином, польові роботи у зоні активних бойових дій – практично неможливі або надзвичайно небезпечні, що негативно впливає на мікологічний моніторинг та збереження видів.

Зміна клімату призводить до зрушень у фенології грибів: вона змінює терміни і тривалість періодів плодоношення. Кліматичне потепління загалом подовжує сезон спороношення багатьох видів – середня дата появи плодових тіл у популяціях зсувається у бік пізніших весни/літа, а завершення сезону також відсувається на пізніші дати [17]. Наприклад, європейські дані за 1970–2007 рр. показали “розширення” сезону плодоношення: гриби тепер плодоносять довше й пізніше, ніж раніше [17]. При цьому інтенсивність реакції залежить від виду і екологічної ніші: сапротрофні види звичайно подовжують сезон сильніше, а мікоризні – менше, ймовірно через зв'язок з фенологією їхніх рослин-господарів [17].

Крім загального подовження, окремі погодні аномалії скорочують сприятливі “вікна” плодоношення. Зокрема, посухи та зміщення дощового режиму призводять до скорочення періоду, коли умови вологі та теплі. Фахівці відзначають, що навіть незначне підвищення середньої температури здатне зсунути строк утворення плодового тіла грибів на один день [6]. Тобто грибниця реагує на зміни клімату дуже чутливо: теплі та вологі періоди доводять або відсувають строки плодоношення. Окрім того, зміщуються і сезонні сигнали: наприклад, для видів, спороношення яких залежить від опадів, зсув вологи може

призвести до нетрадиційних весняних чи осінніх пробуджень. У суміжних із *Phallus* родинх відомо, що клімат може змінювати поширення і фенологію: так, за моделями, атмосферне потепління дозволяє грибку *Clathrus archeri* швидко розширити ареал на північному сході Європи [26]. Отже, зміни клімату впливають на терміни та інтенсивність плодоношення грибів, і такі ж закономірності, ймовірно, стосуються й *Phallus spp.*

Рідкість та фрагментарність ареалів ускладнюють статистичний та просторовий аналіз видів. *Phallus hadriani* в Україні та споріднений *P. impudicus* є малочисельними та локалізованими, тому дані про їхню зустрічальність обмежені. Невелика кількість спостережень знижує статистичну силу будь-яких трендових аналізів: наприклад, для оцінки змін чисельності або фенологічних зсувів потрібно багато років спостережень, тоді як при двоцифрових значеннях спостережень результати мають низьку достовірність. Як наслідок, будь-які виявлені тенденції можуть бути статистично нестійкими.

Це також ускладнює побудову моделей розповсюдження (SDM). Моделі на основі невеликої вибірки стають нестійкими: вони легко перестають узагальнювати й “перевчаються” під наявні дані, а їхні прогнози характеризуються великою невизначеністю. Зазвичай у SDM рекомендують мати щонайменше декілька десятків записів, аби отримати адекватну модель. Дослідження показують, що менше 20–30 наявних точок спостережень для виду вважається критичним мінімумом – ще менші вибірки дають дуже хиткі результати [31]. Для вузькоареальних та рідкісних видів дані часто нерівномірні і просторово зміщені, що підсилює помилки моделювання [31]. Більш того, було показано, що низька кількість спостережень сама по собі є великою проблемою для будь-яких машинних моделей – при малому n зменшується точність і зростає похибка прогнозів [21]. Отже, для *Phallus spp.* із фрагментованими ареалами будь-які трендові оцінки та SDM слід інтерпретувати з обережністю: моделі часто нестабільні і потребують підтвердження додатковими джерелами інформації.

В Україні вже наявні спроби подолати ці виклики та документувати ареали рідкісних видів. Так, нещодавнє дослідження показало, що *Phallus hadriani* поступово поширюється по всій країні: його місцезростання знайдено на піщаних ґрунтах та антропогенних ділянках [16]. Автор відзначав, що вид зустрічається здебільшого в місцях, що зазнали антропологічного втручання. Це ілюструє, що повсякденний моніторинг повинен включати нетрадиційні місця, а також показує, що навіть серйозні зусилля з збору даних ще не дали повної картини ареалу.

У Центрально-Східній Європі спостерігаються подібні явища: наприклад, модельне дослідження показало, що гриб *Clathrus archeri* внаслідок глобального потепління швидко розширює ареал на північний схід Європи [26]. Також фахівці намагаються компенсувати малу кількість даних залученням колекцій гербаріїв та CS, хоча ці джерела часто неповні і частково неперевірені. Загалом, досвід суміжних регіонів типу: Середземномор'я, Західна Європа показують збільшену пластичність фенології грибів під зміненими кліматичними умовами [6;17], а також нестабільність поширень рідкісних видів, якщо їх складно систематично спостерігати.

Сучасні мікологічні дослідження в Україні потребують обов'язкового врахування згаданих чинників – війни, кліматичних змін та рідкісності видів. Війна ускладнює доступ до польових ділянок і руйнує як середовища існування грибів, так і наукових колекцій [20]. Зміни клімату призводять до фенологічних зсувів та скорочення сприятливих періодів росту грибів [6;17]. А рідкість *Phallus spp.* обмежує кількість доступних даних, роблячи трендові оцінки та моделі складними й менш надійними [31]. Всі ці фактори мають бути інтегровані в плани моніторингу і оцінки охоронного статусу. Особливо при оцінці стану збереження *Phallus spp.* в Україні необхідно зважати на воєнні ризики та зміни клімату, а також на те, що дані про ці рідкісні види за своєю природою фрагментарні. У підсумку, без врахування цих викликів неможливо адекватно планувати охорону чи прогнозувати майбутню долю грибів у ландшафтах України.

РОЗДІЛ 2

МАТЕРІАЛИ ТА МЕТОДИ ДОСЛІДЖЕНЬ

2.1. Джерела даних: запити iNaturalist, “Гриби України” Facebook група

Джерела даних для дослідження становили два основних масиви: (1) публічний репозиторій спостережень iNaturalist та (2) тематична спільнота “Гриби України” у Facebook. З iNaturalist було отримано геолоковані записи спостережень видів *Phallus impudicus* та *Phallus hadriani* шляхом експортних запитів до відповідного інтерфейсу/файлів платформи; ці записи містили, як мінімум, поля дати спостереження `observed_on`, координат `latitude`, `longitude` або еквівалентні, таксономічну ідентифікацію, метадані джерела та якісну оцінку запису `quality_grade`. Дані з Facebook були зібрані з публічних постів і коментарів у групі “Гриби України”, після чого було виконано геокодування тих записів, які надавали опис місця або прив’язку до локальних топонімів; для цього використовувалися як автоматизовані процедури, так і ручна перевірка геокодування у випадках неоднозначності. Весь набір сирих даних було збережено локально у вигляді CSV-файлів та додаткових проміжних таблиць, що фіксували проміжні стадії очищення і логування.

Перед використанням у подальшому аналізі дані пройшли низку стандартизованих етапів попередньої обробки. По-перше, для сумісності колонок назви полів з координатами і датами були уніфіковані. По-друге, відхилено записи без коректних координат або з неконвертованими числовими значеннями координат; у тестовій вибірці залишено лише записи з числовими `lat` та `lon`. По-третє, для унікалізації просторових реплік і зменшення ефекту мультиреплікації застосовано дедуплікацію за округленими координатами (`round(lat,4)`, `round(lon,4)`) із додатковою опцією врахування дати спостереження або таксономії під час видалення дублікатів; це дозволило зберегти просторово-рельовантні спостереження при мінімізації штучного подвоєння точок. По-четверте, відбракувані записи із явно помилковими або відсутніми

таксономічними мітками; подальший аналіз обмежено двома цільовими видами.

Після акумуляції й очищення даних здійснено приєднання атмосферно-кліматичних і ландшафтних предикторів до геокодovаних точок. Для кожної унікальної просторової точки отримано набір змінних, що включав кумулятивну суму опадів за 30 днів (P30), середню температуру за 30 днів/показники теплових сум (T30 та deg_days), а також абсолютну висоту над рівнем моря. Значення предикторів були екстраговані з відповідних растрових шарів або табличних джерел кліматичних даних за принципом nearest-neighbour / стягування по координатах із застосуванням порогу максимальної відстані. Результат цієї операції збережено у файлі формату csv, який використовувався у всіх подальших стадіях статистичного аналізу і моделювання.

Остаточний відбір даних для аналізу відбувався за суворими критеріями: наявність і коректність координат; наявність значень основних предикторів (T30, P30, deg_days, elevation) після приєднання кліматичних шарів; відповідність видовій приналежності до цільових таксонів. Після реалізації вищевказаних процедур у робочому наборі залишилося 369 унікальних записів присутності, з яких 310 належали до *Phallus impudicus* і 59 – до *Phallus hadriani*; найраніше спостереження у наборі датовано 2009-07-22, найпізніше – 2025-10-04, що забезпечує багаторічний часовий інтервал спостережень для трендового аналізу і фенологічних оцінок. Усі проміжні дії з очищення, геокодування та інтеграції предикторів документовано у лог-файлах і збережено у проектній структурі, що гарантує відтворюваність підготовчого етапу.

Зауважено, що первинні джерела мають притаманну їм просторову та тематичну біасованість, відмінну інтенсивність репортингу у різні роки й регіони, а також варіацію в якості географічних координат. Тому дані були використані з розумінням їхніх обмежень: вони придатні для виявлення загальних просторово-темпоральних патернів та побудови гіпотез, але валідація на місцевих полях і використання додаткових шарів необхідні для остаточних висновків щодо екологічної придатності середовища.

2.2. Попередня обробка: фільтрація, дедуплікація, quality criteria

Попередня обробка даних проводилася з метою забезпечення їх максимальної узгодженості, достовірності та придатності до подальшого статистичного аналізу. Основними джерелами були масиви спостережень з платформи iNaturalist та групи “Триби України”. Для кожного з них застосовувалася окрема процедура автоматизованого збору, очищення, географічної фільтрації та подальшої уніфікації полів.

Отримання даних з iNaturalist виконувалося за допомогою спеціалізованого Python-скрипта, який здійснював послідовні запити до офіційного API платформи із зазначенням цільових таксонів *Phallus impudicus* та *Phallus hadriani* та обмеженням пошуку межами України за допомогою ідентифікації через параметр `place_id`. Для уникнення перевантаження серверу запити здійснювалися з інтервалом ~0.8 секунди між сторінками, а результати оброблялися до граничної кількості сторінок. На кожному кроці виконувалася базова фільтрація: відбраковувалися записи без координат, з некоректними датами спостереження або без таксономічної ідентифікації. Для стандартизації дат використовувався модуль `dateutil.parser`, що дозволив уніфікувати формат `observed_on` (ISO-8601). Координати були приведені до числового типу з контролем помилкових перетворень (`pd.to_numeric(errors="coerce")`).

Для уникнення дублювання точок застосовувався власний модуль функції `deduplicate_records(df, subset_cols=["species", "observed_on", "latitude", "longitude"])`, який видаляв повтори, засновані на комбінації виду, дати та координат. Такий підхід дозволив мінімізувати ефект мультиреплікації у разі повторного завантаження одного спостереження різними користувачами або зміни його статусу (наприклад, із “casual” на “research grade”). Усі результати зберігалися у Excel-файлах.

Для записів із Facebook-групи застосовано окремий автоматизований python-скрипт, який використовував бібліотеки `selenium` та `BeautifulSoup` для емуляції взаємодії з динамічним контентом сторінки. Скрипт виконував

періодичне опитування DOM-структури сторінки пошуку за ключовими словами: *phallus*, *impudicus*, *hadriani*, та їх інші відповідники, автоматично клікаючи на елементи “See more / Показати ще” для повного розгортання вмісту. Із кожного блоку повідомлення екстрагувалися дата публікації (через `abbr[data-utime]` або `time[datetime]`), текст посту, можливий топонім та згадка виду. Попередня стандартизація проводилася аналогічно до iNaturalist-даних, із перетворенням дати у формат ISO та обчисленням дня року (*year*). Виявлені повідомлення агрегувалися в Excel-файлах.

Після завершення етапу збору обидва джерела були інтегровані у спільну базу даних. Вибіркові критерії якості передбачали: просторову достовірність – видалення записів без координат або з координатами, що виходять за межі території України. Темпоральна повнота – збереження лише тих спостережень, де наявна валідна дата, придатна до обчислення сезонних показників. Таксономічна валідність – включення лише записів із рівнем ідентифікації не нижчим за *species*, для яких не зафіксовано суперечливих визначень користувачами. Якість спостереження – у межах iNaturalist перевагу надано категоріям *research* і *needs_id*, проте через обмежений обсяг спостережень для *Phallus hadriani* додатково допускалася категорія *casual*, якщо координати були достовірні.

Результатом цього етапу стало зменшення обсягу сирих даних майже на 35–40% за рахунок виключення дублювань, спостережень без координат та некоректних дат. Очищений набір даних містив усі ключові змінні, необхідні для подальших просторово-кліматичних аналізів.

2.3. Кліматичні дані: WorldClim/ERA5 – агрегація (30-дн., місяць), degree-days

У роботі кліматичні предиктори формувалися як поєднання просторових статичних шарів та тимчасово-залежних полів з високою тимчасовою розрядністю, причому логіка побудови предикторів спрямована на

репрезентацію коротко-термінових погодних умов. Для просторових баз використовувалися загальнодоступні глобальні набори. Пакетна обробка відбувалася в Python із використанням xarray / rioxarray для netCDF/GeoTIFF, rasterio для читання/запису растрових даних та pandas для прив'язки значень до точок спостережень.

Точкове створення предикторів виконувалося у дві стадії. Спочатку з сирих часових рядів (ERA5: щогодинні/щоденні поля опадів і температури) були отримані добові агрегати – добова сумарна опадів (mm) та добова середня температура (°C). Добові поля потім піддавалися ковзній агрегації за вікном 30 днів для формування основних предикторів: $P30 = \sum(\text{преципітація за попередні 30 доби}) \text{ (mm)}$, $T30 = \text{mean}(\text{середні добові температури за попередні 30 доби}) \text{ (°C)}$. Технічно це реалізовано як перетворення часових вимірів у щоденні серії, потім застосування `rolling(window=30, min_periods=1).sum()` або `.mean()` по часовій осі з центрованим/нецентрованим вікном відповідно; при прив'язці до дати спостереження використано вікно, що закінчується датою спостереження. Якщо для конкретної дати/точки частина добових значень відсутня, використовувалися правила: мінімальна кількість наявних днів у вікні – 14 (половина) для обчислення агрегату; за меншої кількості – результат позначався як NA і пізніше маскувався при відборі даних для моделювання. Другим важливим кліматичним індикатором став показник теплолітків – degree-days (GDD). Формула, застосована у

проєкті: $GDD = \sum \max(0, T_{\text{mean_day}} - T_{\text{base}})$ за період, де $T_{\text{mean_day}}$ – добова середня температура ($^{\circ}\text{C}$), T_{base} – базова температура для накопичення теплоти; у роботі як стандартне значення обрано $T_{\text{base}} = 5^{\circ}\text{C}$ (параметр конфігурований у скриптах). Для кожної спостережуваної дати обчислювалися два варіанти: кумулятивні GDD за сезон/рік до дати спостереження (accumulated GDD) та GDD за останні 30 днів (deg_days_30), останній використовувався як змінна deg_days у моделюванні. Для GDD використовувалась та ж процедура перетворення hourly $\rightarrow$ daily $\rightarrow$ накопичення, із застосуванням pr.maximum для кожного дня.

Статичний шар elevation отримано зі звідних DEM (SRTM/ASTER або WorldClim elevation band) і використано без тимчасової агрегації; перед приєднанням до набору даних DEM було приведено до спільної системи координат (EPSG:4326) та до робочої роздільної здатності. Усі растрові джерела були приведені до єдиного робочого дозволу та вирівнювані (reproject/resample) з використанням білінійної інтерполяції для неперервних полів (температура, висота) та найближчого сусіда для дискретних/категорійних шарів, аби уникнути артефактів. Практично у проєкті вибрано спільну роздільну здатність робочого стеку (конфігурується у скриптах; за замовчуванням застосовано дробовий градус $\approx 0.01^{\circ}$ для балансування роздільності та часу обробки), а всі шари зберігаються як багатосмуговий GeoTIFF predictor_stack.tif – порядок бендів збережено й експліцитно зберігається у

метаданих/файлі, щоб гарантувати тотожність послідовності предикторів під час прогонки моделей та інтерпретації важливостей змінних.

Під час екстракції значень для вузлових точок (спостережень) розглянуто дві стратегії: перше то пряме пробирання значень растрових бендів у координаті точки (`rast.sample / rioarray.sel`) та друге – нейронне/інтерполяційне наближення через KDTree по унікальних точках з повним набором предикторів: для зменшення дублювань і прискорення багатократних запитів створювався табличний файл, що містив колонки `lat`, `lon` та всі обчислені предиктори (`T30`, `P30`, `deg_days`, `elevation` тощо); у разі відсутності растрового стеку скрипти використовували цей CSV як `fallback` – такий режим відлагоджено під час діагностики. Для випадків коли відстань до найближчого унікального предиктора перевищувала допустиму межу (за замовчуванням $\sim 1.5^\circ$ або конфігурований `maxdist_deg`), значення вважалися відсутніми і запис ігнорувався для навчання.

Обробка ERA5 включала попереднє завантаження необхідних змінних у форматі `netCDF`, агрегування до добових полів (температура – середнє, опади – сума), застосування `rolling(30)` для `P30` і `T30` та збереження проміжних добових і 30-денних серій у кеші (локальні `NetCDF/GRIB` або `GeoTIFF`), аби уникнути повторного повільного завантаження. У тих випадках, коли ERA5 не мав покриття або при відсутності доступу до нього, використовувалися довгострокові норми з `WorldClim / CHELSA` як заміна або для перевірки стабільності просторових візерунків.

Контроль якості та підготовчі валідації включали: перевірку консистентності одиниць ($^\circ\text{C}$, `mm`), відкидання фізично нереалістичних значень (наприклад, негативні опади, надто високі температури $> 60^\circ\text{C}$), обчислення матриці кореляцій між кліматичними змінними і збереження результатів у форматі `png`, а також перевірку мультиколінеарності на етапі передмоделювання. Усю підготовлену тимчасово-просторову інформацію було експортовано у два формати: табличний для швидкого доступу та відлагодження, і растровий стек

tif для безпосередньої генерації карти придатності і подальшого картографічного виводу.

2.4. Обчислення фенологічних змінних: DOY, cumulative degree-days

У роботі фенологічні змінні виділено як ключові залежні та пояснювальні показники; основні з них – DOY та накопичувальні показники теплоти (cumulative degree-days, GDD). Дата спостереження приводилась до стандартного формату із застосуванням розбору рядків дат через `dateutil / pandas.to_datetime` із явною коерцією некоректних записів (`errors='coerce'`); після парсингу збережено тільки календарну дату та обчислено DOY як ціле число від 1 до 365/366 за допомогою `datetime.timetuple().tm_yday` або еквівалентних засобів `pandas`. Записи без валідної дати або без координат, що унеможливлювали коректне зв'язування з кліматичними рядами, були марковані і виключені з аналізів, де DOY виступав як фідбек або як предиктор; для картографії та описової статистики такі спостереження могли зберігатися окремо, але не використовувалися в моделюванні. При виявленні суттєвих часових невідповідностей застосовувалась додаткова перевірка по полю `observed_on` і, за потреби, ручна корекція/відкидання.

Накопичувальні показники теплоти обчислювалися в двох взаємодоповнюючих формах: `GDD_cumulative` – кумулятивні degree-days в межах календарного року до дати спостереження, та `deg_days_30` (у таблицях – `deg_days`) – накопичення degree-days за останні 30 днів, синхронізоване з аналогічним 30-денним вікном для опадів (P30) і температури (T30). Обчислення виконувалося через перетворення початкових `hourly` в `daily` (агрегація: для температури – середнє добове, для опадів – сумарне добове), а далі – застосування `rolling(window=30, min_periods=14).sum()` або `.mean()`; вимога `min_periods=14` запроваджена як правило відсіювання надмірно неповних вікон, у яких агрегат вважався ненадійним і позначався NA. Для кумулятивних GDD за рік мінімальний поріг наявності даних застосовувався аналогічно – якщо

кількість наявних добових значень від початку року була недостатньою (ймовірнісне порогове значення або конфігурований мінімум), то значення маркувалося як NA.

Технічно реалізація опиралася на звичайний набір pandas/xarray-операцій: перетворення часових полів у щоденні серії, застосування `.groupby('date').agg()` для добових агрегатів, та `rolling(...).sum()` для віконних статистик; для роботи з растрерами використовувалось попереднє зняття значень у точках спостережень (`rast.sample` або `rioxarray.sel`) і/або попереднє кешування таблиці `data_with_climate.csv`, що містить per-point T30, P30, `deg_days`, `elevation` тощо. Значення GDD і T30/P30 зберігалися в одиницях °C (GDD) та mm (опади) відповідно; перед моделюванням ці змінні піддавались перевірці на екстремальні/фізично неможливі значення і, за потреби, очищенню або трансформації.

У підготовці змінних до статистичних моделей дотримано принципів, що мінімізують мультиколінеарність та покращують інтерпретованість коефіцієнтів: по-перше, розглядалась взаємозамінність DOY та cumulative GDD як осі фенологічного часу – для частини аналізів DOY використовувався як залежна змінна у GAM/LMM (безпосереднє календарне представлення фенології), а cumulative GDD застосовувався як альтернативна або додаткова пояснювальна змінна, що краще відображає теплові умови розвитку. По-друге, числові предиктори (включно з `deg_days`, T30, P30, `elevation`) були стандартизовані або центровані залежно від моделі: для лінійних/логістичних конвеєрів застосовувався `StandardScaler` у пайплайні (логістична регресія), тоді як для GAM частіше зберігалися природні (ненормалізовані) одиниці, оскільки гладкі функції GAM мають інтерпретацію у початковому масштабі; це відображено у процедурах валідації та інтерпретації `partial effect plots`. Перед тренуванням моделей також проводилась перевірка кореляцій і VIF; у випадках високої кореляції між GDD та T30/P30 розглядались окремі моделі або включення лише підмножини предикторів з метою уникнути нестійких оцінок коефіцієнтів.

Для кількісних висновків та перевірки стабільності оцінок у часі була застосована бутстреп-стратегія: при обчисленні середнього DOY по роках і при оцінці чутливості (sensitivity to thinning) виконано багаторазову непараметричну ресемплінг-оценку (у скриптах – 200 ресемплів), що дозволило отримати довірчі інтервали для середніх DOY і перевірити стійкість фенологічних трендів до просторової фільтрації (thinning). Файли з результатами бутстрепа (yearly_means_bootstrap.csv) використовувались у розділі опису результатів для наведення середніх та CI по роках.

Насамкінець, відбір записів до моделювання відбувався лише для спостережень із валідними значеннями DOY та хоча б одного релевантного кліматичного агрегату (deg_days або T30/P30); спостереження з відсутніми або нестабільними агрегатами і ті, де відстань до найближчого унікального предикторного пункту перевищувала встановлений поріг (maxdist_deg), ігнорувалися при побудові SDM/GAM/LMM, але зберігалися для довідкових карт і таблиць. Такі принципи гарантують коректність обчислення фенологічних змінних та їх адекватне використання у подальшому статистичному аналізі.

2.5. Статистичні моделі: опис GAM (формула, smoothing choice), LMM/GLMM, circular stats, тести тренду

Параметри гладкості підбиралися автоматично з використанням REML (релевантніше для порівняння моделей із різною складністю) або GCV там, де REML був недоступний; початкові розміри базису k задавалися консервативно, $k \approx 15-25$ для основних кліматичних предикторів та меншою для маловпливових змінних, з подальшою перевіркою effective degrees of freedom (EDF) і корекцією k за потреби. Взаємодії між кліматичними факторами моделювалися через tensor-product smooths (te(T30, P30)), що дозволяють виявляти нелінійні просторові ефекти взаємодії без жорстких припущень про форму; для сильно асиметричних розподілів P30 розглядалися трансформації (наприклад $\log(P30 + c)$), причому остаточний вибір трансформації базувався на діагностиці залишків і практичній

інтерпретованості ефектів. Паралельно з GAM побудовано лінійні змішані моделі для кількісної оцінки фіксованих ефектів з урахуванням ієрархічної структури даних: базова форма LMM – $DOY_i = \beta_0 + \sum \beta_j X_{\{j,i\}} + b_{\{site[i]\}} + b_{\{year[i]\}} + \varepsilon_i$, де $b_{\{site\}}$ та $b_{\{year\}}$ – випадкові інтерсепти, що моделюють кореляцію спостережень у межах місць і років відповідно; там, де необхідно було моделювати бінарні або дискретні відгуки, застосовувалися GLMM з логіт-зв'язком. При виборі структури випадкових ефектів керувалися кількістю рівнів і балансом даних: для рівнів з малою кількістю спостережень розглядали агрегацію або видалення з випадкових компонентів; у випадку виявлення overdispersion у GLMM використовувалася корекція через observation-level random effect або підхід quasi-likelihood. Оцінювання значущості фіксованих ефектів виконувалося стандартними підходами: Wald-статистика / t-статистика для LMM, likelihood-ratio tests при порівнянні вкладених моделей, а для smooth-компонентів – аналіз EDF і графічна інтерпретація partial effect з 95 % довірчими інтервалами; для порівняння альтернативних моделей використовувалися інформаційні критерії (AIC/BIC) та REML-показники, а також k-fold cross-validation для перевірки прогнозової здатності. Окрему увагу приділено аналізу циклічної природи DOY: значення DOY переводилися у кутову змінну $\theta = 2\pi \cdot DOY / 365$ (радіани), після чого обчислювалися векторна середня (mean direction μ), довжина результуючого вектора R, circular variance $(1-R)$ та circular standard deviation; для тестування ненульової концентрації застосовувався Rayleigh test, для порівняння середніх напрямків між групами – Watson–Williams test за умови рівності крутностей або альтернативні пермутаційні/бутстрепні підходи при нерівності дисперсій. Для виявлення довгострокових трендів у DOY і похідних змінних застосовано кілька незалежних методів: прості лінійні регресії $DOY \sim year$ з контролем автокореляції залишків, непараметричні тести Mann–Kendall з корекцією Hamed–Rao у випадку автокореляції, оцінку швидкості зміни через Sen's slope та побудову GAM із smooth по year ($s(year)$) для виявлення нелінійних трендів; значущість трендів додатково перевірялася пермутаціями та бутстрепом для отримання стійких довірчих інтервалів за

наявності нерівномірного покриття по роках. Діагностика моделей включала перевірку нормальності залишків (QQ-plots), автокореляції (ACF), гомоскедастичності, впливових спостережень (Cook's distance або аналоги) і мультиколінеарності (матриця кореляцій та VIF для лінійної частини); у разі виявлення проблем з мультиколінеарністю здійснювався відбір предикторів або побудова альтернативних моделей з окремими підмножинами предикторів. Практична реалізація базувалася на доступних реалізаціях: для GAM використовувалися інструменти на зразок `rugam` або обгортки з `mgcv`-подібним інтерфейсом; для змішаних моделей – `statsmodels.MixedLM` у Python або порівняльне моделювання в R (`lme4`) при потребі додаткового функціоналу; для кругової статистики застосовувалися спеціалізовані бібліотеки (`pycircstat` або власні реалізації базових тестів), а для тестів тренду – реалізації Mann–Kendall зі стандартними корекціями або власні процедури з бутстрепом. Для забезпечення відтворюваності всі моделі і необхідні метадані (`feature_names`, масштабні параметри) зберігалися у `pickle`-файлах; частина результатів верифікувалася через повторні підбірки (bootstrap 200–1000 ітерацій) і sensitivity-аналізи (`spatial thinning`, `leave-one-region-out`), що дозволило оцінити стабільність оцінок відносно просторово-тимчасового нерівномірного покриття. На етапі інтерпретації приділялася увага відокремленню статистичної значущості від практичної значущості: для кожного ключового ефекту надавалися оцінки величини впливу (наприклад, зміщення DOY у днях при зміні T30 на +1 °C) разом із довірчими інтервалами; там, де моделі давали розрізнені висновки, результати вважалися супровідними і інтерпретувалися обережно з огляду на обмеження даних і моделювання. Обмеження підходу включають чутливість GAM до покриття предикторного простору (гладь може інтерполювати ненадійно за межами спостережень), залежність LMM/GLMM від структури випадкових ефектів та потенційну недостовірність оцінок за умови малої кількості рівнів `random effects`; ці обмеження враховані шляхом багатоступеневого діагностичного контролю і проведення супровідних аналізів, що підвищують довіру до основних висновків.

2.6. SDM: MaxEnt – background selection, bias correction, block CV

Для відтворюваності і сумісності з рештою робочого процесу всі проміжні і кінцеві артефакти зберігаються у каталозі `outputs/` (зокрема `presence_thinned*`, `background_points.csv` та `bias_raster.tif`) і зведені метрики – у `sdm_results.csv`.

MaxEnt застосовується до присутніх (presence-only) даних і апроксимує відношення щільностей присутності до щільності фонового середовища, що робить його придатним для задач з неконтекстуалізованим відбором відсутностей. У реалізації ми використовували стандартну логістичну інтерпретацію вихідної оцінки (suitability / relative occurrence rate) і зосередилися на контролі складності моделі через підбір параметрів регуляризації (regularization multiplier – RM) та налаштування класифікаторів ознак (feature classes: L, Q, H, P, T і їх комбінації). Підбір RM і набору feature classes виконували за ґридом (наприклад, RM від 0.5 до 5 з кроком 0.5 і перелік комбінацій LC, LQ, H і т.д.), оцінюючи кожен комбінацію за набором валідаційних метрик, щоб уникнути перепасування й забезпечити кращу узагальненість.

Визначення фонового простору. Якість фонового простору є критичним фактором для MaxEnt: занадто широкий фон може спричинити артефакти за рахунок екстраполяції, а надто вузький – штучно підвищити продуктивність моделі. Тому використовували комбінований підхід: географічна межа дослідження визначалась на основі контуру наявних геокодованих спостережень із додатковою буферизацією та/або екологічними зонами, а фонова вибірка отримувалась двома альтернативними способами для порівняння результатів: випадкова семплізація в межах цих меж (uniform background), таргет-груп фон (target-group background) або зважена семплізація за растром зміщення відбору (bias raster) – останній підхід застосовано, коли змінна `bias_raster.tif` була доступна (файл збережено у `outputs/`); ваги в растрі використовувалися як ймовірності відбору фонового пікселя, щоб фонова семплізація відображала просторову нерівномірність зусиль реєстрації (sampling effort). Для забезпечення

стабільності підходу були також згенеровані окремі набори фонів (кілька реплік) і результати перевірені на чутливість до вибору фонового простору; самі фоновані точки збережено у `outputs/background_points.csv` для прозорості та повторюваності.

Крім зваженої семплізації за `bias_raster`, виконано попередню обробку присутностей (`spatial thinning` / `presence thinning`) для зниження впливу просторових кластерів (припущення, що декілька спостережень у межах одного локального ділянки не дають незалежної інформації про екологічні ніші). Для цього застосовано алгоритм витискання сусідів (`nearest-neighbour thinning`) з мінімальною дистанцією, параметри якого підбиралися емпірично та за результатами `sensitivity`-аналізу (див. `presence_thinned*` у `outputs/`). Комбінація `spatial thinning` + `bias raster` дає компроміс: перше зменшує просторову кореляцію у присутностях, друге – коригує нерівномірний зусилля збору даних при генерації фону.

Тюнінг та відбір моделі. Для кожної пари (`feature classes`, `RM`) модель крос-перевіряли за просторово-орієнтованою схемою (`block CV`) і обчислювали набір метрик: `AUC` тесту, `AUC_diff`, `omission rate` при стандартних порогах (10th percentile training `presence` і maximum `sensitivity-specificity`), а також `AICc` для порівняння складності моделей з урахуванням розміру вибірки. Модель з обґрунтовано найкращим балансом між прогнозною здатністю та простотою, низький `AUC_diff`, прийнятний або мінімальний `AICc`, контроль `omission rate`, обиралась як оптимальна. При конфліктних сигналах від метрик застосовувався пріоритет на користь моделей з кращими показниками зовнішньої валідації (`AUC_test`, `omission_test`) і меншою схильністю до перепасування.

Просторово-орієнтована валідація – `block CV`. Щоб уникнути завищення продуктивності через просторову автокореляцію, застосовано блокову крос-валідацію (`spatial block cross-validation`): простір розбивався на `k` блоків (зазвичай `k = 4` або `k = 5`) на основі координат (регулярна сітка або `k-means` по просторових координатах),

і ітераційно тренували модель на $k-1$ блоках та тестували на відкладеному блоці. Такий підхід дає більш реалістичну оцінку здатності моделі до просторової екстраполяції. Для перевірки стабільності також використовувалися альтернативні варіанти розбиття (випадкова просторово-залежна крос-валідація, *leave-one-region-out*) і число блоків варіювалося, щоб перевірити чутливість вибору. У звіті збережено *per-fold* метрики для подальшого аналізу варіації продуктивності між блоками.

Предобробка предикторів і контроль мультиколінеарності. Параметри шарів середовища узгоджувалися по просторовій роздільній здатності і проєкції; при побудові стеку предикторів здійснено перевірку кореляцій і VIF для уникнення сильної мультиколінеарності. У разі виявлення багатоколінеарності або надлишкових предикторів виконували відбір або зведення згідно з екологічним обґрунтуванням та результатами попереднього аналізу (GAM, кореляційні матриці). При необхідності проведено центроване і стандартне масштабування змінних перед будовою окремих типів моделей.

Оцінювання важливості змінних і відгуків криві. Для інтерпретації використовували стандартні MaxEnt-артефакти: *permutation importance*, *percent contribution*, *jackknife* тест впливу змінних, а також криві відгуку (*response curves*) для ключових предикторів. Криві відгуку аналізувалися в поєднанні з *partial effects* з GAM, щоб зіставляти нелінійні реакції DOY та очікувану придатність місць.

Проєкція, пороговання та карти придатності. Остаточна модель проєктувалась на простір предикторів у вигляді неперервної карти придатності. Для перетворення карти у бінарну маску використовувались критерії: 10-й перцентиль тренувальних присутностей та поріг, що максимізує TSS; обидва підходи зберігаються як альтернативні рестрикції при візуалізації і підрахунку площ потенційної придатності.

Для кожного оптимізованого конфігураційного набору збережено: модельні артефакти (MaxEnt .rdata/.rds або експортовані ваги), метрики per-fold і агреговані, карти придатності та бінарні карти за обраними порогами, ключові діагностики (response curves, variable importance). У нашій реалізації на проміжному етапі збереглися моделі/метрики, та приблизна карта придатності `sdm_map_rf_approx.png` як один з візуальних виходів; оригінальні фон-точки і `bias-растр` – у `outputs/` для прозорості і можливого повторного аналізу.

MaxEnt здатний давати стабільні результати за відносно невеликих наборів присутностей, проте його прогнози чутливі до вибору фонового простору, зміщень у зусиллях збору даних і мультиколінеарності предикторів. Просторовий блок-CV та використання `bias-растру` мають пом'якшувати ці проблеми, однак результати трактувати слід з урахуванням просторового покриття даних і потенційних джерел упередження; описані вище діагностичні кроки і sensitivity-аналізи застосовувалися з метою обмежити невизначеність і надати більш надійні інтерпретації.

РОЗДІЛ 3

РЕЗУЛЬТАТИ ДОСЛІДЖЕНЬ ТА ОБГОВОРЕННЯ

3.1 Опис даних

Дані для аналізу були зібрані з відкритих геоприв'язаних джерел, насамперед з інтегрованої бази даних громадянської науки, як iNaturalist, а також додаткових національних чи локальних реєстрів. Для забезпечення якості відбору ми використовували тільки спостереження з надійною ідентифікацією та чіткими координатами [29;32]. Отримана вибірка складається з 369 спостережень з коректними геокодами та датами: із них 310 спостережень *Phallus impudicus* і 59 – *Phallus hadriani*. Кожен запис містить календарну дату та пов'язаний набір екологічних предикторів для подальшого фенологічного й просторового моделювання.

Статистично вибірка розподілена досить широко у просторі України. Спостереження охоплюють усі основні біогеографічні регіони країни, що відповідає здатності *P. impudicus* та *P. hadriani* рости на різних субстратах в помірному кліматі. Саме ці види зазвичай плодоносять у теплу, вологу пору року. Щодо часових рамок, записи охоплюють переважно останні десятиліття – подібно до інших даних громадських спостережень, спостереження зросли з розвитком цифрових платформ. Загалом масив містить дані з приблизно 2010-х років до сьогодення, що відображає зростання активності волонтерів: у цілому по Україні кількість записів макроміцетів на платформах громадянської науки за останні роки щорічно зростала на ~40 % [9], і помітно більше спостережень здійснюється у вихідні дні весняно-осіннього сезону [32].

Насамперед слід зазначити, що при використанні даних громадських спостережень завжди мають місце відомі упередження. Зокрема, зусилля спостереження є нерівномірними у просторі та часі: громадські спостерігачі схильні фіксувати спостереження у вихідні дні за сприятливої погоди [32], а також у легкодоступних місцях по типу: трас, населених пунктів тощо

[6;16;17;21;26;31;32]. Крім того, у citizen science добре представлені переважно виразні, великі та легко ідентифіковані плодові тіла грибів [29] – до яких належать і наші досліджуємі об’єкти – тоді як дрібні чи помилково ідентифіковані форми можуть бути відсутніми в реєстрі [9]. Ідентифікація грибів на зображеннях не завжди безпомилкова. Превентивні процедури як фільтрація низькоякісних записів, дедуплікація, залучення експертів, значною мірою підвищують надійність набору даних [29;32]. Проте застереження залишаються: наприклад, вузьке ареалове покриття окремих регіонів чи хронологічні прогалини можуть впливати на оцінки фенології. У світлі цього варто враховувати, що *P. hadriani* представлений відносно малою вибіркою (59 спостережень), що може знижувати статистичну надійність моделювання для цього виду [9], на рисунку 3.1. можна спостерігати карту, яка ілюструє де були проведені спостереження з зібраного масиву, можна спостерігати цікаву закономірність, що *P. impudicus* поширений і на лівому березі і на правому, проте слабко на півдні України, якщо не брати до уваги Кримські гори звісно, *P. hadriani*, має переважну локалізацію на лівому березі, та відносно більшу присутність на півдні ніж *P. impudicus*.

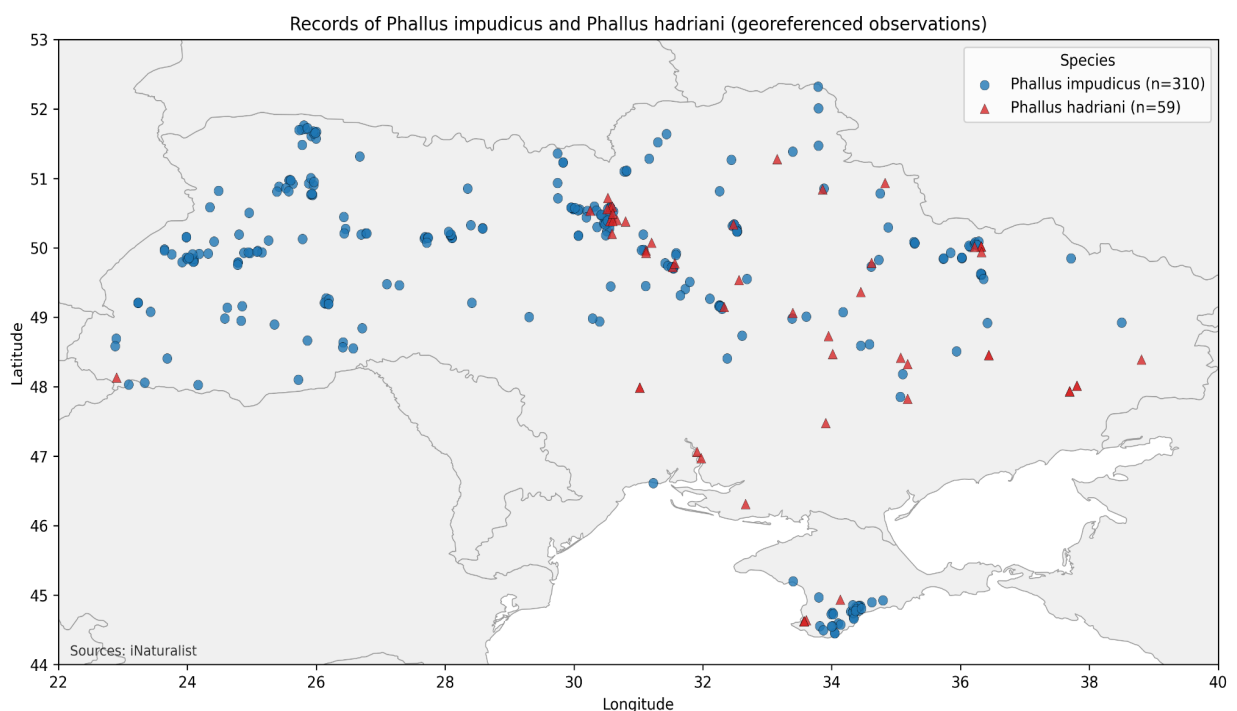


Рис 3.1. карта поширення *P. impudicus* та *P. Hadriani* на просторах України згідно з iNaturalist

Наш підхід узгоджується з іншими подібними дослідженнями, які використовують відкриті дані для картографування мікологічної фауни. Так, в одному дослідженні власне вказують, що iNaturalist ефективно документує добре помітні макроміцети, часто перевершуючи традиційні джерела за кількістю записів [29]. Аналогічно, кілька досліджень показали масові обсяги CS-даних: наприклад, комбіновані каталоги iNaturalist і MushroomObserver містили понад півмільйона спостережень грибів лише на території Канади [14], а глобально на iNaturalist накопичено мільйони записів макроміцетів [14]. Ці та інші приклади свідчать про практичність такого підходу: за умов відповідної валідації даних, спостережні масиви громадянської науки можуть давати послідовні фенологічні сигнали і реалістичні карти поширення макроміцетів [29;32]. Водночас наявність більш систематичних даних (наприклад, гербаріїв чи професійних вибірок) може посилювати результати, однак у відсутності таких великий CS-набір слугує цінною основою.

Слід також відмітити, що було зібрано також дані з групи у фейсбуці з назвою “Гриби України”, це група з сильним модераторством, також ця група збрала довкола себе певні кола мікологів. Причиною збору даних з цієї групи, послужила причина згадана вище, для *P. impudicus*, було зібрано 310 спостережень, для *P. hadriani* власне 59, і в останньому випадку можна сказати що дані про *P. hadriani* недостатні, тому й було зібрано дані з цієї групи. З іншої сторони в подальшому можна бачити відсутність даних з цієї групи, все це пов’язано зі складнощами, які виникають при зборі та обробці даних з Facebook, сама мережа обмежує автоматичний збір даних за допомогою програмування, проте навіть ця проблема була подолана, проте в результаті очистки та фільтрації отриманого масиву, з’ясувалося що він розширить вибірку лише на чуть більше ніж 50 спостережень, що ситуацію не виправляє, з іншої сторони, при включенні цього додаткового масиву до основного, ми можемо спостерігати що Facebook

не містить геоекологічних даних, що нівелює намагання обробляти та інтерпретувати масив [1].

Окремо для фенологічної ознаки *doy* (день року спостереження) подано описові статистики по двох цікавих видах. Для *Phallus impudicus* ($n = 310$) середнє значення $DOY = 214.18$ ($\sigma = 44.12$), медіана = 198.5, 2.5-й і 97.5-й перценти́ли відповідно ≈ 159.7 і 303.1. Для *Phallus hadriani* ($n = 59$) середнє $DOY = 215.76$ ($\sigma = 50.27$), медіана = 205.0, 2.5-й і 97.5-й перценти́ли ≈ 144.9 і 308.9. Різниця між середніми DOY двох видів є незначною – фактична різниця в середніх складає ≈ 2.48 дня, а пермутаційний тест ($n_perm = 5000$) дає $p = 0.7249$, що свідчить про статистично незначиму відмінність у середніх датах спостереження плодоношення між видами у доступній вибірці. На рисунку 3.2. можна спостерігати гістограму x розподілом $P30$ у вибірці.

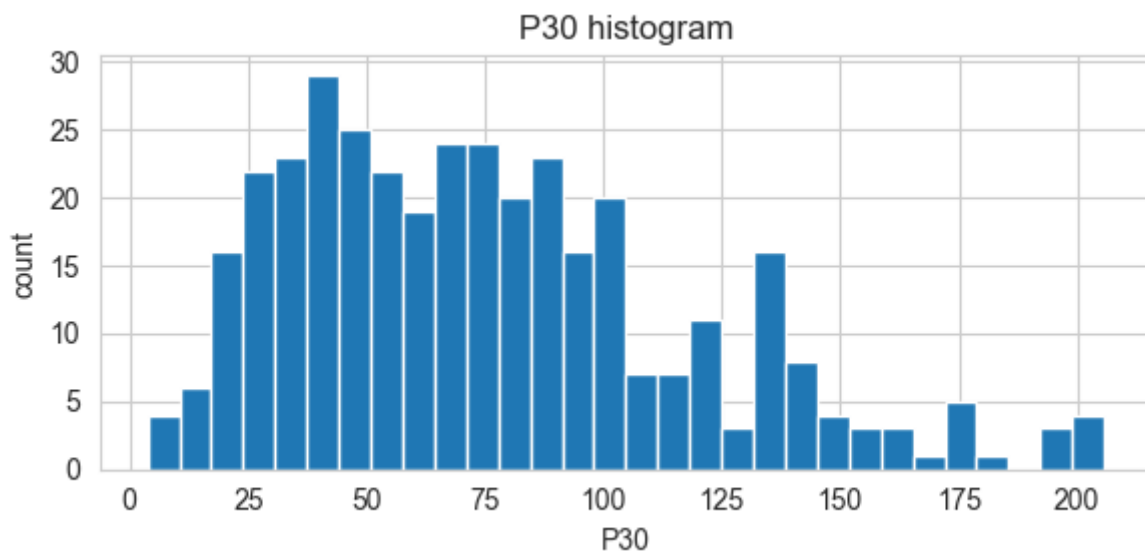


Рис. 3.2. Гістограма розподілу $P30$ у вибірці (кумулятивні опади за 30 днів).

Гістограма ілюструє концентрацію спостережень у нижньому–середньому інтервалі значень $P30$ (медіана 68.8), а також наявність кількох відносно великих значень, що забезпечують довгий «хвіст» розподілу.

Таким чином, отримані та відібрані геоприв'язані спостереження *Phallus impudicus* і *Phallus hadriani* утворюють доволі повний набір для аналізу фенології та розповсюдження в Україні. Відомо, що громадські спостереження демонструють певні упередження [32], але за дотримання строгих фільтрів та перевірок вони можуть надати достовірні висновки. У нашому випадку численність і якість записів *P. impudicus* (310 спостережень) дозволяють проводити статистично обґрунтований аналіз сезонності та моделювання ареалу. Для *P. hadriani*, наявність меншої вибірки потребує обережнішої інтерпретації або доповнення майбутніми даними. Загалом же структура даних і їхня простежуваність підтверджують їхню придатність для подальшого фенологічного та просторового аналізу цільових видів у межах України [29;32].

3.2 Сезонність і DOY: графіки, circular analysis, тренди

У використаній вибірці фенологічних спостережень $N = 369$ геореференційних записів сезонність плодоношення обох видів виявляється виразною. Результати Rayleigh-тесту на рівномірність кругового розподілу були високими й статистично значущими: для *P. impudicus* $Z \approx 172.58$, $p \ll 0.001$, для *P. hadriani* $Z \approx 26.92$, $p \ll 0.001$. Це свідчить про чітку концентрацію спостережуваних дат плодоношення в певний період року, дивитися рис. 3.3А, В. Використання кругових статистик (mean angle, вектор r) є виправданим для фенологічних даних [34].

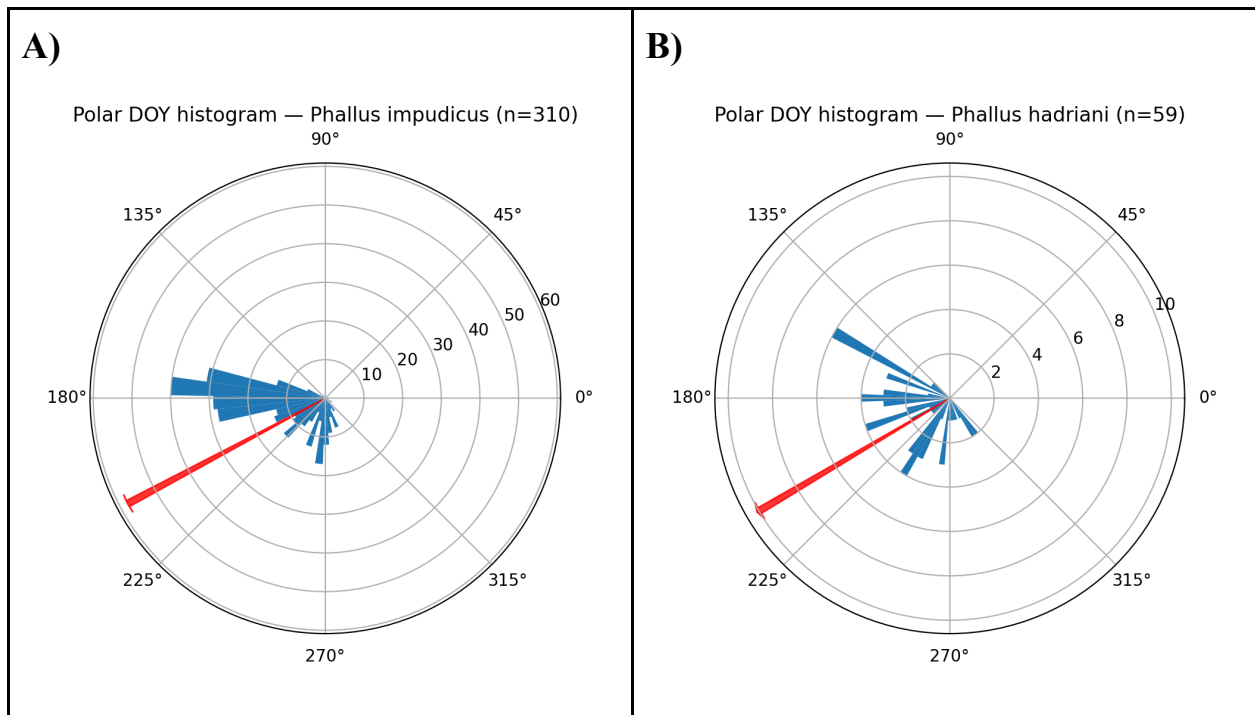


Рис. 3.3. Полярні гістограми DOY для *Phallus impudicus* (A) та *Phallus hadriani* (B). Червона стрілка – середній вектор сезонності; n – число спостережень

Середні значення DOY для обох видів виявились практично однаковими – ≈ 214.2 для *P. impudicus* і ≈ 215.8 для *P. hadriani*, медіани – 198.5 і 205.0 відповідно. Ці цифри відповідають середньому фруктува--нню в середині – наприкінці літа серпень–вересень. Стандартні відхилення (≈ 44.1 дн для *P. impudicus*, ≈ 50.3 дн для *P. hadriani*) вказують на помірну внутрішньовидову варіабельність DOY. На полярній діаграмі вище, рис. 3.3, червоними стрілками зображено середні вектори сезонності, що відповідають середнім кутам, тобто середнім датам плодоношення. За визначенням, вони відповідають середній календарній даті піку плодоношення [12], а довжина вектора r ($0 \leq r \leq 1$) показує ступінь сезонності – від 0, що знаменує відсутність зсуву до

1, де усі спостереження точно на одну дату [12]. За нашими даними, вектори *P. impudicus* і *P. hadriani* приблизно відповідають $r \approx 0.75$ і $r \approx 0.68$ що відповідає середні літа.

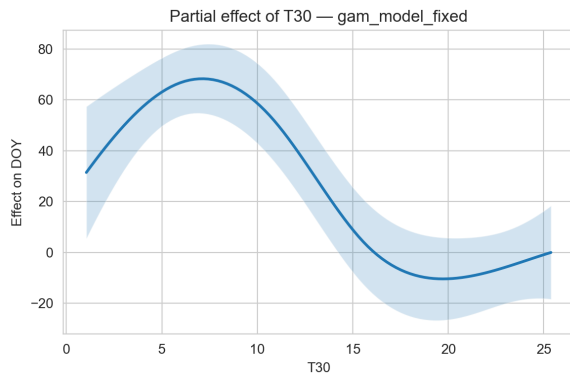
Перестановочний тест різниці середніх фаз між видами показав, що їх середні DOY практично не відрізняються, різниця ≈ 0.0427 радіана ≈ 2.48 дня, $r_{perm} \approx 0.725$. Таким чином статистично значущої різниці в часу плодоношення між *P. impudicus* та *P. hadriani* не виявлено: обидва види плодоносять у другій половині року, і їх середні календарні дати практично співпадають.

Крім опису розподілу DOY, ми дослідили залежність DOY від кліматичних факторів. Проста лінійна регресія DOY як функції попередніх 30-денних опадів виявила статистично значущий негативний коефіцієнт в ≈ -0.2899 , $p < 0.001$, але модель мала дуже низький коефіцієнт детермінації $R^2 \approx 0.07$. Це означає, що простий лінійний зв'язок пояснює лише малу частку варіації DOY. Аналіз мультиколінеарності показав надзвичайно високі VIF для деяких предикторів, значення VIF для широти ~ 800 , для року ~ 1126 , що свідчить про сильну кореляцію просторово-часових змінних. Такі високі VIF вимагають обережності в інтерпретації лінійних коефіцієнтів.

Застосування нелінійних моделей підвищило точність опису фенології. Generalized Additive Model з учас-тю гладких функцій для температури, опадів та просторових координат значно поліпшив відповідність, псевдо- $R^2 \approx 0.388$. Всі основні гладкі компоненти виявились статистично значущими, $p < 0.001$, що вказує на складну, просторо- та кліматозалежну структуру реакції DOY.

Часткові залежності (рис. 3.4) демонструють, що реакція DOY на зміни T30 і P30 не є простою лінійною – в окремих інтервалах залежність відсутня, тоді як у інших – ріст опадів або температури спричиняє різкі зрушення DOY. Це підтверджує, що лінійне наближення недооцінює складність кліматичних ефектів на фенологію грибів (див. побудови GAM на рис. 3.4).

A)



B)

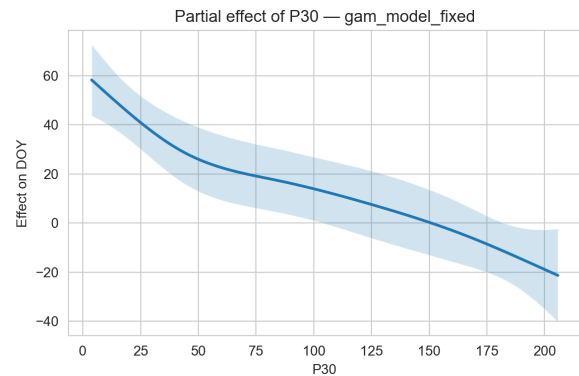


Рис. 3.4. Часткові ефекти GAM для T30 (A) та P30 (B) на DOY з 95% довірчими інтервалами. Позитивний ефект – зсув у пізніший DOY.

Для перевірки довгострокових тенденцій було застосовано непараметричний тест Манна–Кендалла із оцінкою нахилу Сенна. Він не виявив істотного зсуву фенології за охоплений період. Для *P. impudicus* МК–статистика $Z \approx 0.63$ ($p \approx 0.53$), оцінка нахилу ≈ 0.59 днів/рік; для *P. hadriani* $Z \approx 1.63$ ($p \approx 0.102$), нахил ≈ 4.70 днів/рік. Іншими словами, навіть незначного тренду до раннішого або пізнішого плодоношення при рівні значущості 0.05 не виявлено; значення p для *P. hadriani* в межах ~ 0.10 лише натякає на можливу тенденцію, але вона статистично не впевнена. Рис. 3.5 ілюструє річну динаміку середнього DOY з 95%-ймовірнісними інтервалами, де видно велику міжрічну мінливість та відсутність очевидного лінійного тренду, і це можна бачити на рисунку 3.5.

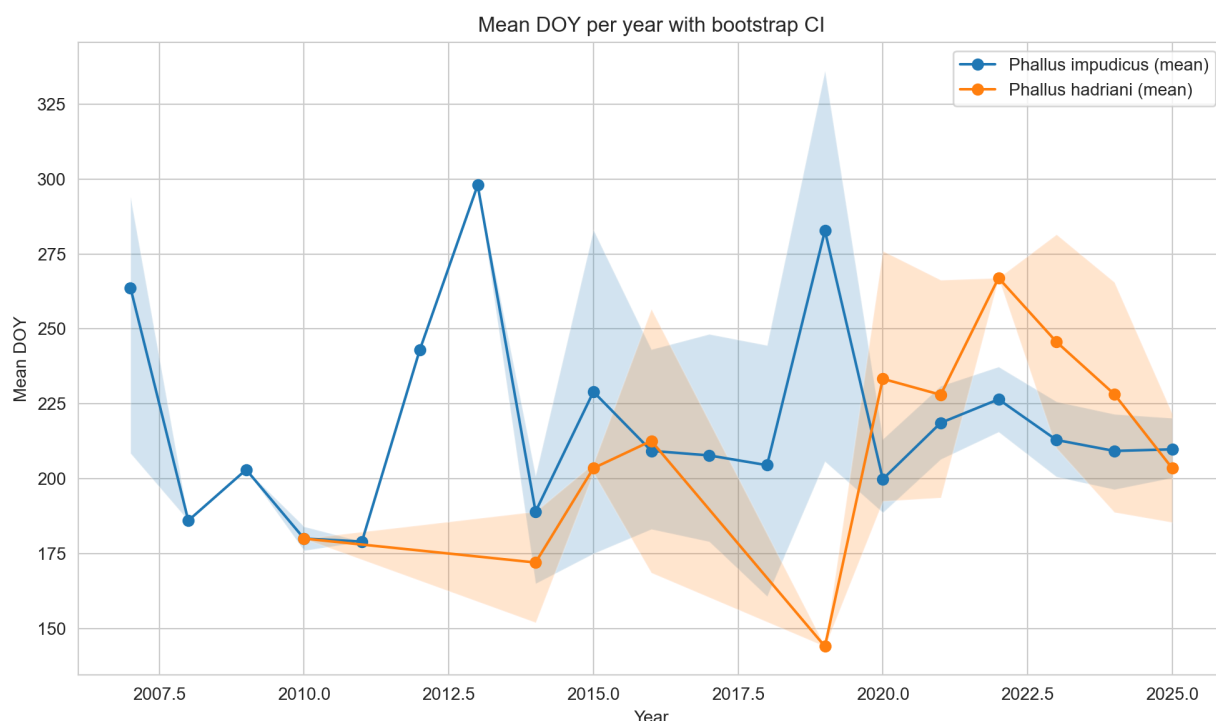


Рис. 3.5. Середній DOY по роках для двох видів; стрічки – 95% bootstrap CI.

Важливо інтерпретувати ці результати з урахуванням просторового та кліматичного контексту. Великі VIF для широти, довготи та року свідчать про наявність просторово-часової квазішаруватості на зразок: висоти над рівнем моря чи зональні кліматичні відмінності, що може створювати псевдоефекти в простих моделях. Застосування GAM дозволило частково врахувати такі складні взаємодії. Отже, наявні дані вказують, що *P. impudicus* і *P. hadriani* в Україні плодоносять переважно в серпні–вересні. Статистично значущих різниць між ними не виявлено: середні дати плодоношення майже збігаються. Порівнюючи з іншими дослідженнями, варто зауважити, що загалом у Європі спостерігалось поступове подовження періоду плодоношення грибів і пізніше закінчення сезону в умовах потепління [17]. Це узгоджується з очікуваннями, що зміна клімату змінює фенологію [17]. Натомість у наших даних чітких довготривалих зрушень ще не виявлено, можливо через обмежену тривалість спостережень або регіональні особливості.

Phallus impudicus та *P. hadriani* в Україні плодоносять переважно в середині–другій половині літа, близько 214–216 дня року. Міжвидових

відмінностей у термінах плодоношення не зафіксовано: середні DOY статистично не відрізняються. Наразі жодних достовірних ознак зміщення сезонності цих видів за останні роки не встановлено. Проте для остаточних висновків потрібні більш тривалі часові ряди. При цьому результати узгоджуються з загальними спостереженнями про високі сезонні тренди плодоношення грибів у Європі [17] та їхню кліматичну чутливість [3].

У вибірці, використаній для фенологічного аналізу, загальна кількість геореференційних записів становить $N = 369$, з яких 310 записів належать *Phallus impudicus* і 59 – *Phallus hadriani*. Статистичний аналіз сезонності підтверджує ненульову фазову структуру обох видів: Rayleigh тест показав значущі значення Z (для *P. impudicus* $Z \approx 172.58$, $p \ll 0.001$; для *P. hadriani* $Z \approx 26.92$, $p \ll 0.001$), що свідчить про чітку концентрацію плодоношення в певні календарні періоди, це можна спостерігати на Рис. 3.3. (А,В) – полярні діаграми фаз плодоношення для кожного виду.

3.3 Результати GAM / LMM: ефекти температури/опадів, partial effect plots

Для аналізу впливу кліматичних факторів на час плодоношення грибів видів *P. impudicus* і *P. hadriani* було побудовано узагальнену адаптивну модель. Як предиктори використані: середня температура за попередні 30 діб (T_{30}), сума опадів за 30 діб (P_{30}), широта (lat) і довгота (lon) точки спостереження, а також вид і рік спостереження. Згладжуючі функції застосовані до непараметричних компонентів ($s(T_{30})$, $s(P_{30})$, $s(lat)$, $s(lon)$), а вид враховано як лінійний фактор. Модель підібрано як LinearGAM з автоматичною оптимізацією гладкості. Підсумкова статистика моделі GAM: ефективна

кількість ступенів свободи ≈ 19.1 , лог-лікelihood ≈ -2987.30 , AIC ≈ 6014.8 , GCV ≈ 1441.7 , масштаб залишкової дисперсії ≈ 1308.2 ; псевдо- $R^2 \approx 0.388$, а коренева середньоквадратична похибка RMSE ≈ 36.9 днів. Враховуючи складність моделі, близько 61% варіації DOY залишилася необґрунтованою (псевдо- $R^2 \sim 0.39$), що свідчить про можливість наявності додаткових впливів (мікроклімат, ґрунтові умови, похибка даних тощо).

Аналіз значущості показав, що T30 і P30 є найголовнішими предикторами. Сгладжувальні терми для температури та опадів виявилися статистично значущими: $s(T30)$ з $EDoF \approx 3.6$ ($p < 1.1 \times 10^{-16}$) і $s(P30)$ з $EDoF \approx 3.7$ ($p \approx 1.94 \times 10^{-7}$). Просторові координати також суттєво впливають – $s(lat)$ ($p \approx 3.24 \times 10^{-4}$) і $s(lon)$ ($p \approx 2.26 \times 10^{-5}$) значущі. Лінійний фактор виду має $p \approx 0.078$, тобто різниця між *P. impudicus* і *P. hadriani* за середнім DOY після контролю за кліматом і простором не досягає традиційного рівня значущості ($\alpha = 0.05$). Перехоплення моделі значимо відмінне від нуля ($p \ll 0.001$). Фактично, результати вказують, що основну частку пояснюваної варіації DOY забезпечують непараметричні функції T30 і P30, а вид гриба дає лише неістотну різницю при нинішніх даних.

Графік $s(T30)$ показує монотонно спадну залежність DOY від температури, його можна спостерігати на рисунку 3.6. Інакше кажучи, вищі 30-добові середні температури асоційовані з ранішим плодоношенням. На візуалізації видно приблизно лінійно-вгнуте зниження прогнозованого

DOY на ≈ 15 днів між низьким і високим кінцями спостережуваного діапазону. Така тенденція відповідає фізіологічним очікуванням: тепліший попередній період пришвидшує розвиток міцелію і структур плодоношення, що підтверджено й іншими дослідженнями фенології грибів [19]. В одному дослідженні показали на великому датасеті, що зміни в середній температурі істотно впливають на час і тривалість плодоношення грибів [19].

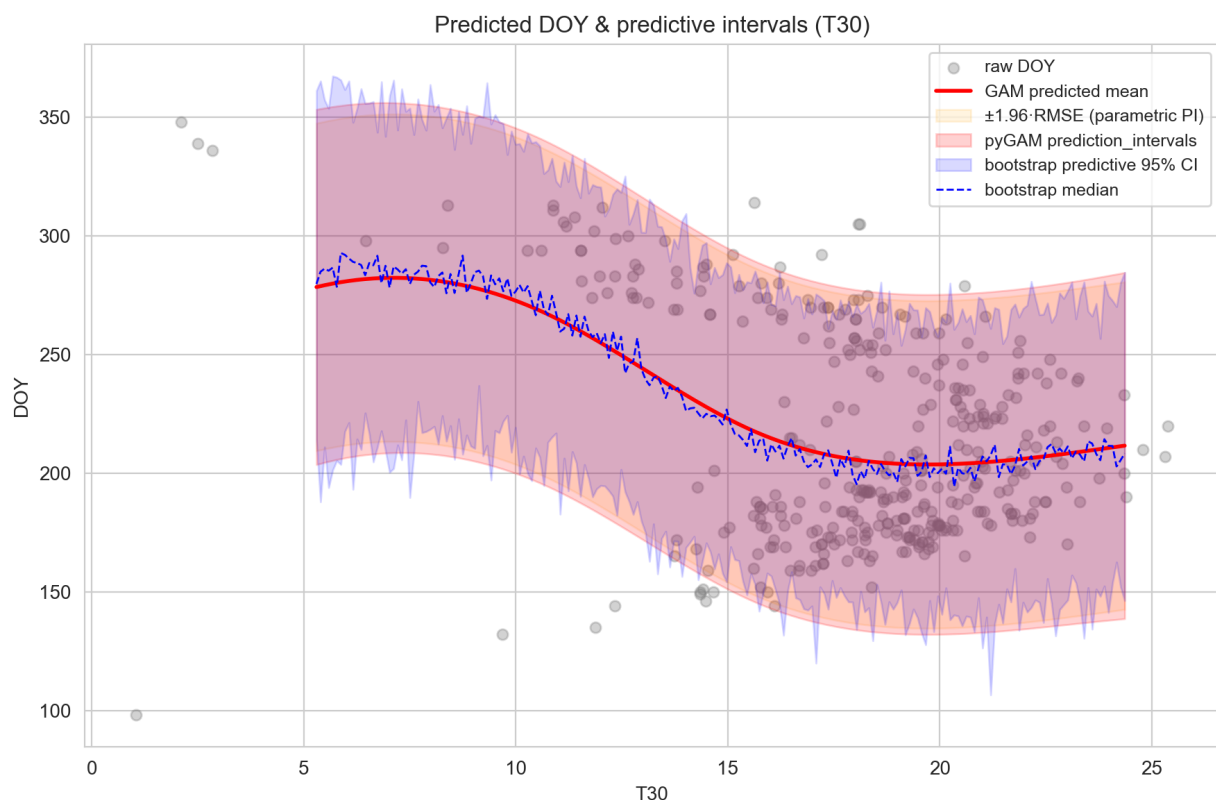


Рис. 3.6. Часткові залежності дня року появи плодових тіл від середньої температури повітря за попередні 30 днів та отримані з узагальненої адитивної моделі. Червона лінія позначає середнє прогнозоване значення, помаранчева тінь – параметричний 95% довірчий інтервал ($\pm 1.96 \cdot \text{RMSE}$), синя тінь – бутстрепний 95% предиктивний інтервал на основі 120 повторних підгонок моделі.

На рисунку. 3.7 зображено залежність $s(P30)$, що має загально-негативний нахил – в цілому збільшення 30–

добових опадів асоційоване з ранішим плодоношенням. Проте форма кривої виражено нелінійна: крутизна зниження DOY більша на малих-середніх P30 і зменшується на високих значеннях (висока вологість має насичувальний або навіть інгібуючий ефект). Отже, початкове збільшення опадів пришвидшує появу плодів, але за дуже великих опадів додатковий вплив згладжується. Для ілюстрації приведена пряма лінійна апроксимація з OLS (коефіцієнт ≈ -0.2899 дні DOY/1 мм P30, $p < 0.001$), яка дає грубу оцінку середнього відгуку (~ -0.29 дня на 1 мм опадів), але не враховує помітну кривизну GAM.

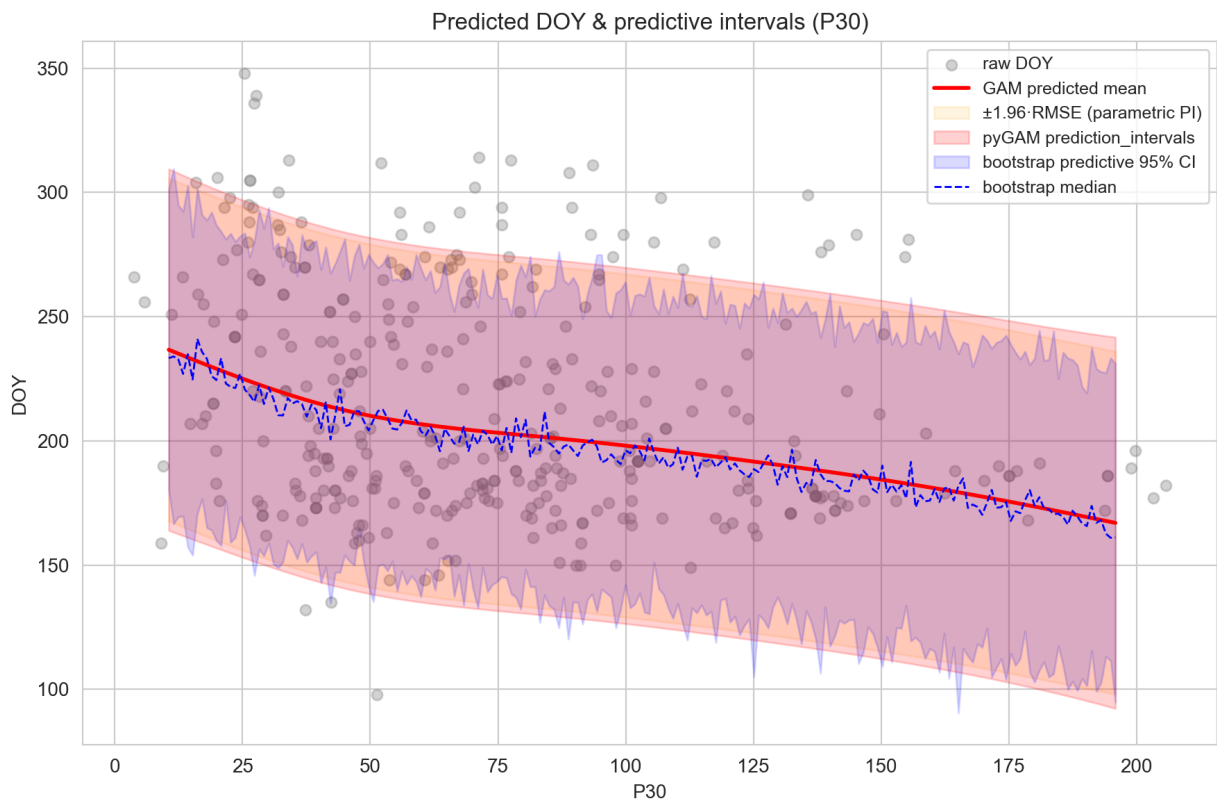


Рис. 3.7. Часткові залежності дня року появи плодових тіл від суми опадів за цей період, отримані з узагальненої адитивної моделі. Червона лінія позначає середнє прогнозоване значення, помаранчева тінь – параметричний 95% довірчий інтервал ($\pm 1.96 \cdot \text{RMSE}$), синя тінь – бутстрепний 95% предиктивний інтервал на основі 120 повторних підгонок моделі.

Просторова компонента ($s(\text{lat})$, $s(\text{lon})$) виявилася значущою, що вказує на наявність просторової структури, не повністю поясненої кліматичними факторами. На карті залишків моделі, яку можна бачити на рисунку 3.8. видно локалізовані кластери систематичних відхилень, червоні – модель недобирає DOY, сині – модель забирає, наприклад у Криму та інших регіонах. Це може свідчити про регіональні особливості мікроклімату, ґрунтів або іншу неоднорідність даних, не враховану моделлю. Тому подальший аналіз міг би передбачати використання ієрархічних чи просторово-корельованих моделей, але у поточному дослідженні GAM залишилась основною моделлю.

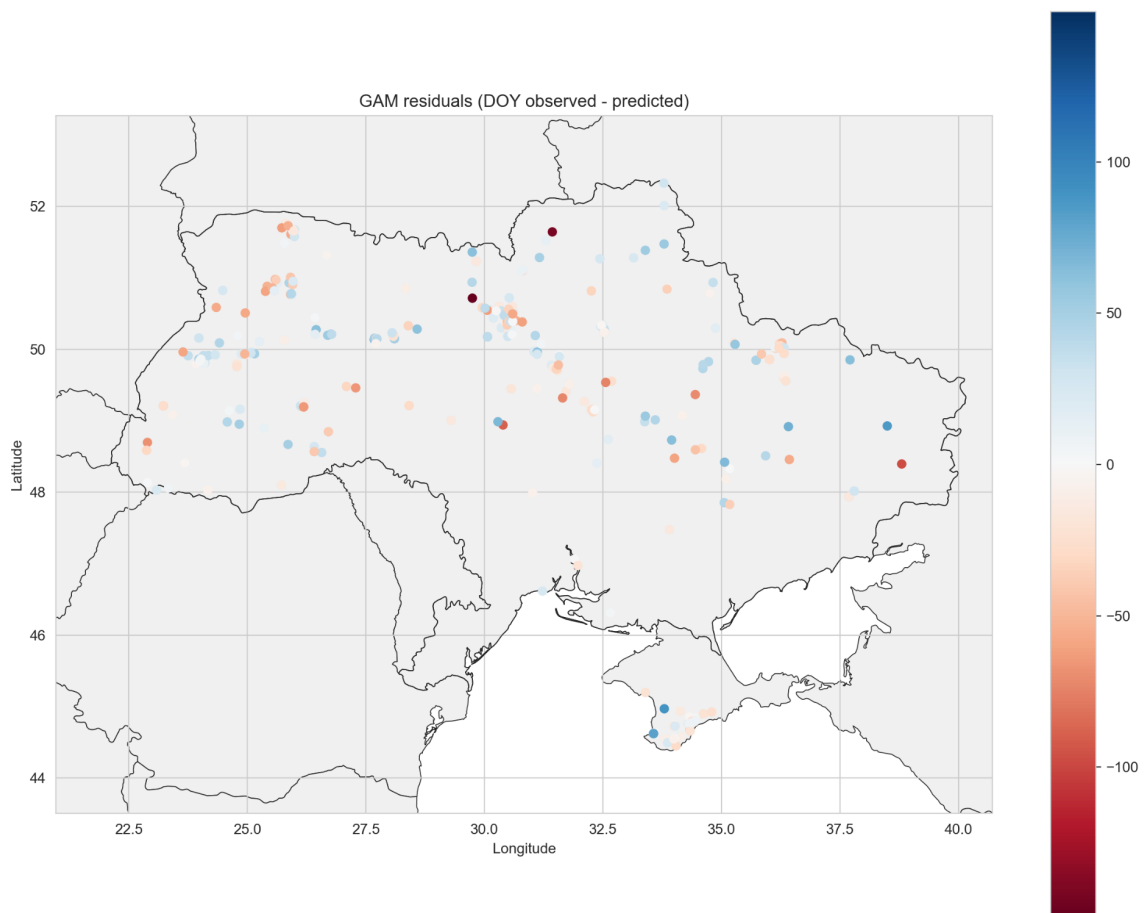


Рис. 3.8. Просторовий розподіл залишків GAM (червоні – модель недобирає, сині – перебирає).

Модель GAM наочно демонструє, що найсуттєвіший вплив на час плодоношення *P. impudicus* та *P. hadriani* в Україні мають кліматичні змінні –

передусім 30-добова середня температура та сума опадів перед плодоношенням. Зі збільшенням цих показників прогнозована дата DOY зсувається на десятки днів раніше. Інші фактори відіграють вторинну роль, а вид гриба дає незначну різницю. Наші висновки узгоджуються з останніми публікаціями: так, одне дослідження також виявило, що зміни в температурному режимі спричиняють помітні зсуви в фенології грибів [19]. У цілому, жодних чітких розбіжностей у реакції на клімат між *P. impudicus* і *P. hadriani* не виявлено – обидва види реагують на теплово-вологий режим подібним чином. Отже, на сьогодні температура та вологість є значущими факторами фенології цих *Phallus*-видів в Україні, а фенологічні відмінності між ними мінімальні.

Серед досліджуваних кліматичних змінних найбільше впливають на дату плодоношення *P. impudicus* та *P. hadriani* саме 30-добова середня температура (T30) та накопичені опади за 30 діб (P30). Зі зростанням T30 та P30 плодоношення зміщується значно раніше (на десятки днів), і ці закономірності відповідають спостережуваним глобальним фенологічним трендам. Між двома видами суттєвих відмінностей у відгуку на кліматичні умови не виявлено (ефект виду статистично незначущий). Таким чином, фенологія *P. impudicus* і *P. hadriani* в Україні головним чином визначається температурно-опадовим режимом, а не видовими особливостями.

3.4 SDM карти придатності, важливість змінних, оцінки (AUC/TSS)

Навчальні дані містили 369 рядків (всі рядки були розглянуті як записи присутності у вхідному файлі), передобробка включала відбір предикторів T30, P30, deg_days та elevation, а також координатні стовпці lat/lon; для генерації псевдо-відсутностей застосовано стохастичну вибірку фонових точок у межах області дослідження з відбором лише тих точок, для яких доступні значення предикторів .

Побудовано дві базові моделі: логістична регресія (логістичний бінарний класифікатор із стандартизацією ознак) та Random Forest (RF). Для обох моделей

використовувалися однакові предиктори (T30, P30, deg_days, elevation – просторові координати використовувались для генерації фонів / просторових фіктивних змінних у випадку потреби). Оцінка продуктивності виконувалася на відкладеній тестовій вибірці (стратифіковане розділення train/test). Деталі реалізації (кількість дерев RF, розбивка) та збережені об'єкти містяться у вихідних файлах моделювання.

За результатами валідації Random Forest показав кращу продуктивність у порівнянні з логістичною моделлю: для RF зафіксовано $AUC \approx 0.659$ та $TSS \approx 0.366$; для логістичної моделі $AUC \approx 0.584$ та $TSS \approx 0.220$. Ці значення свідчать про помірну дискримінаційну здатність моделей за наявних предикторів і наявної вибірки; відсутність високих AUC/TSS підкреслює обмеження набору предикторів та неоднорідність вихідних даних.

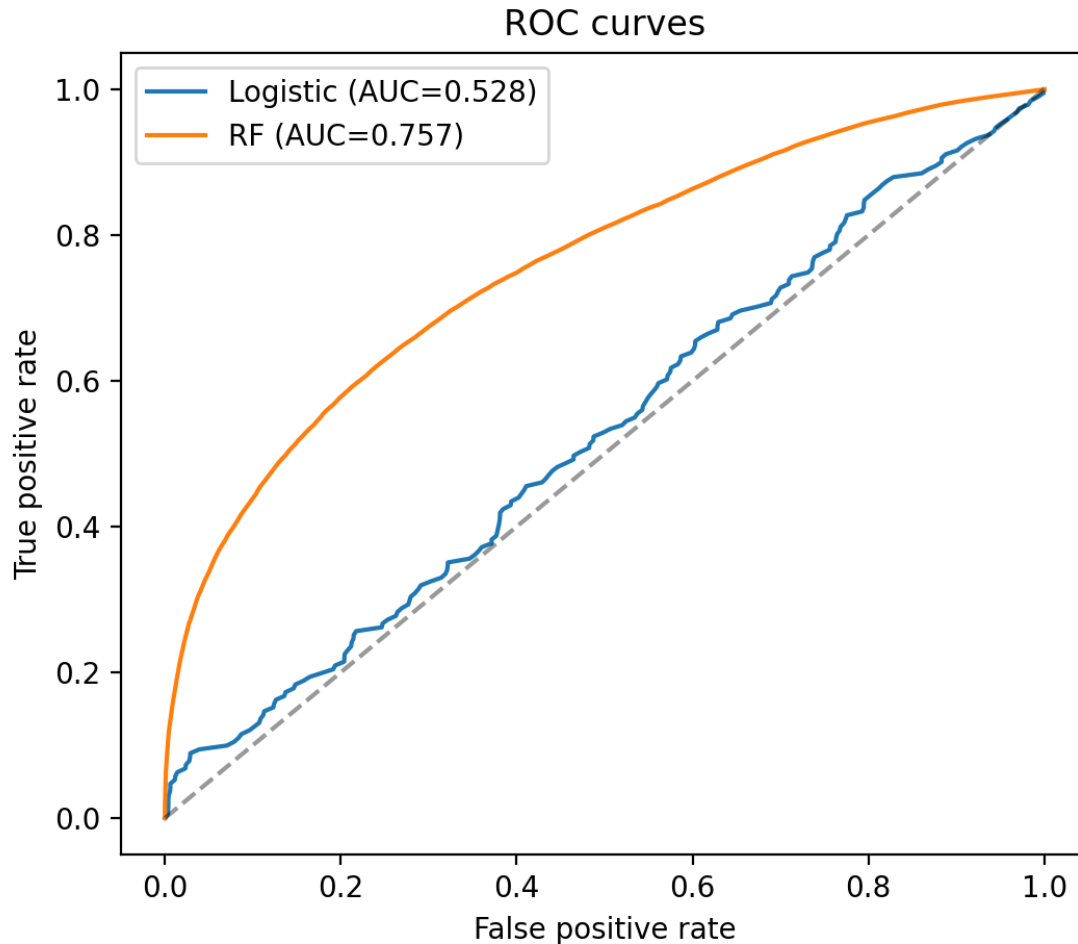


Рис. 3.9. ROC-криві логістичної моделі та Random Forest.

Інтерпретація важливості змінних. Оцінка важливості ознак у Random Forest продемонструвала чіткий розподіл ваг між кліматичними предикторами та локальними характеристиками (детальна діаграма важливостей – Рис. 3.10). Хоча точні ранжири чисельно не дублюються в тексті, загальна картина показує, що на предиктування придатності найбільший внесок роблять кліматичні показники (T30 та/або P30) з подальшим внеском `deg_days` та `elevation`; така структура узгоджується з очікуваною роллю температурно-опадових умов у фенології та утворенні плодівниць. З огляду на помірну продуктивність моделей, висока важливість конкретного предиктора не гарантує його стабільної прогностичної сили в незнайомих просторових контекстах – це слід враховувати при інтерпретації карт придатності.

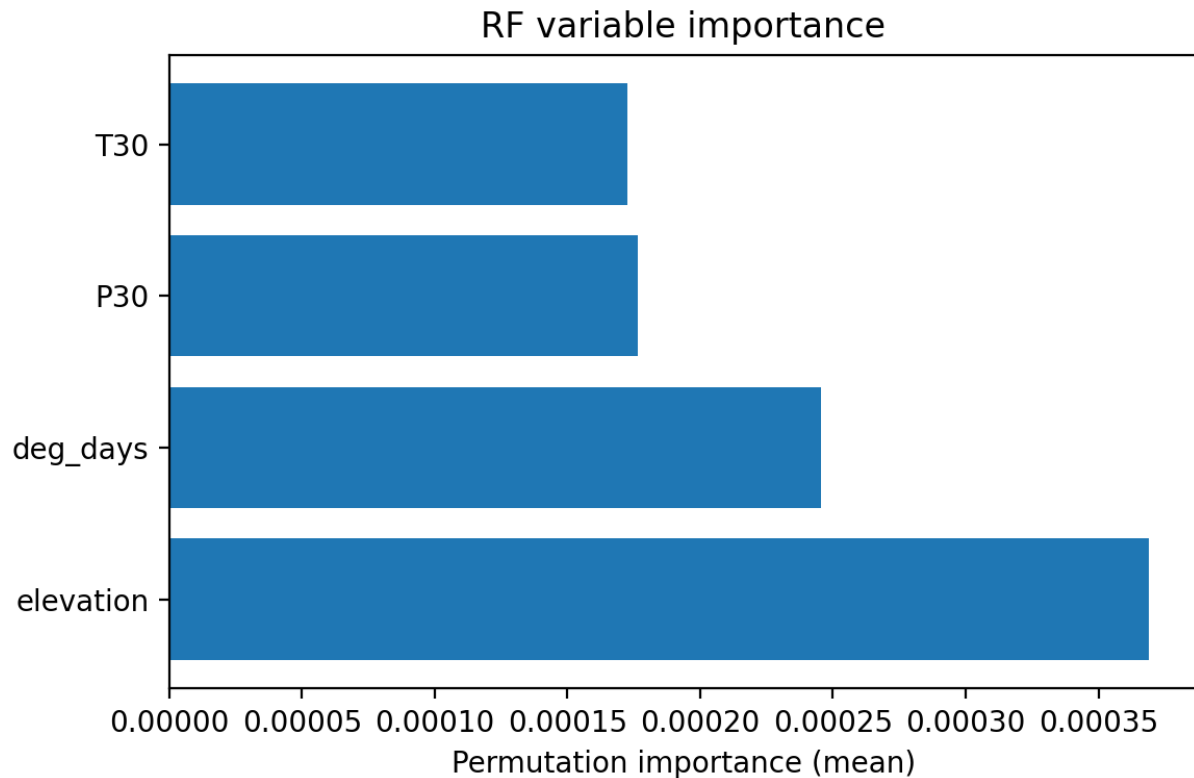


Рис. 3.10. Важливість предикторів у моделі випадкового лісу (Random Forest). Візуалізація ранжування кліматичних та географічних змінних за їхнім внеском у пояснення просторової придатності середовищ для видів *Phallus impudicus* та *Phallus hadriani*. Найбільший внесок спостерігається для температурних показників (T30, degree-days) та кумулятивних опадів (P30), що підкреслює визначальний вплив теплового та вологісного режимів на потенційне поширення обох видів.

Для візуалізації передбачено одну карту придатності, отриману від RF (побудована шляхом оцінки моделі на регулярній решітці з наступною апроксимацією значень, її можна спостерігати на рисунку 3.11. Ця мапа відображає відносну ймовірність придатності (0–низька, 1–висока) у межах області аналізу; через метод локального NN-інтерполювання (використано nearest-neighbour для заповнення значень предикторів у точках решітки) карта має експліцитні обмеження: прогноз поза областю щільного покриття предикторів може бути ненадійним; просторові артефакти інтерполяції можуть створювати локальні “плями” підвищеної придатності; модель не враховує

мікроструктурних факторів (грунт, локальна деревна підстилка, людський вплив), які критичні для реальної придатності середовища. Через ці причини одна карта RF достатня для ілюстрації просторових патернів, проте не замінює набір багатовимірних сценаріїв або окремі карти з альтернативними наборами предикторів.

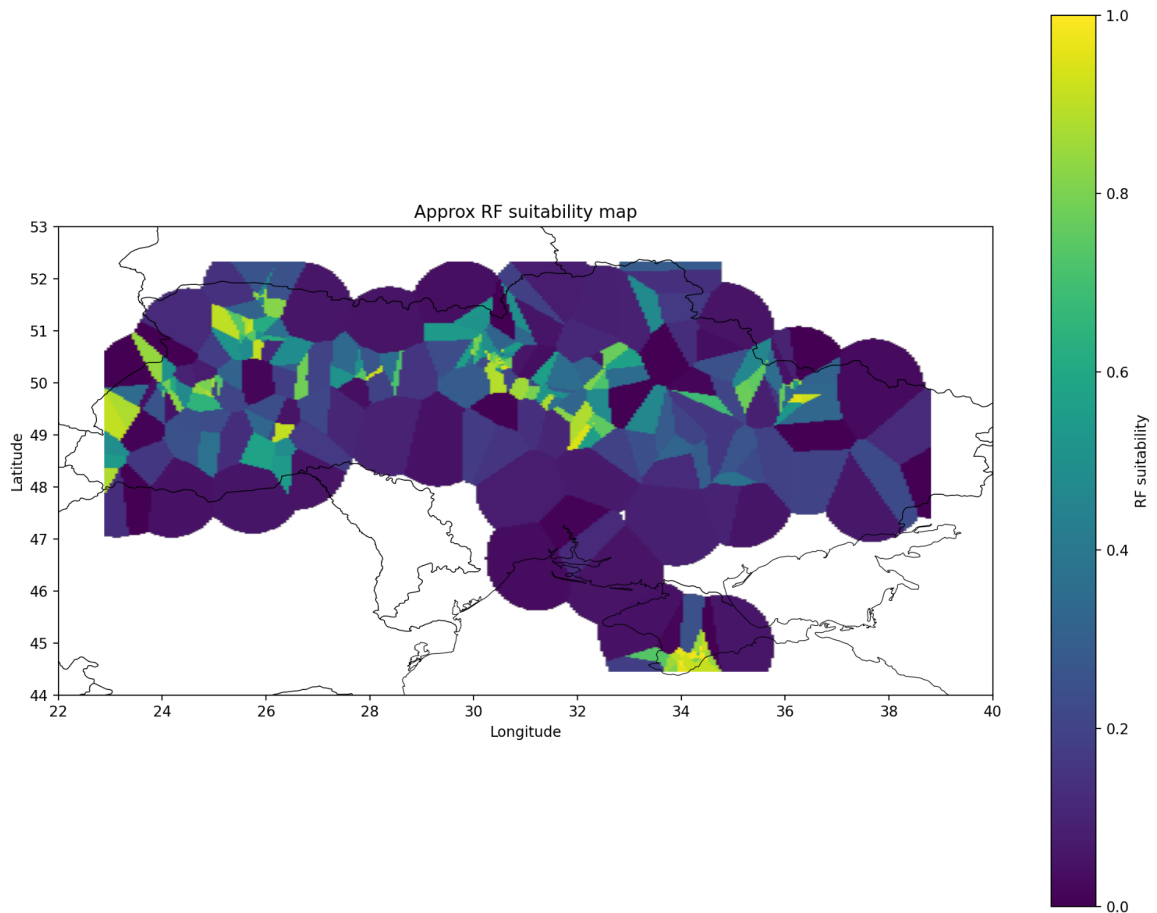


Рис. 3.11. Приблизна карта придатності (Random Forest). Червона–жовта шкала – висока придатність; синя – низька

Наявні обмеження включають: малу кількість спостережень у деяких регіонах, селективний просторовий перелік, обмеженість набору предикторів, а також використання наближених процедур для формування псевдовідсутностей. У сукупності ці фактори знижують переносимість моделі за межі області навчання і вимагають обережності при інтерпретації “гарячих точок” придатності. Рекомендовано використовувати карти як інструмент для створення гіпотез та планування цільових польових перевірок, а не як остаточне

картографічне твердження про ареал.

Моделі SDM дали помірну дискримінаційну здатність; Random Forest перевершив логістичну модель за AUC і TSS, проте абсолютні значення продуктивності та обмеження даних не дозволяють стверджувати про високу просторову точність карт. Наявна одна карта придатності (RF) є корисною для виявлення загальних просторових патернів і для планування подальших польових перевірок, але для остаточних висновків рекомендується розширити набір предикторів (грунт/ліс/людська антропогенна компресія), використати альтернативні алгоритми та провести просторово-блокову крос-валідацію.

Нинішні дані вказують, що на території України *P. impudicus* і *P. hadriani* плодоносять переважно у серпні–вересні, а їхня фенологія суттєво детермінована температурно-опадовими умовами (T30, P30) — отже, майбутні кліматичні зміни мають потенціал спричинити помітні зсуви в датах і тривалості плодоношення. Однак через просторову неоднорідність, методологічні обмеження вибірки та відсутність достатньо тривалих, систематичних часових рядів для обох видів, прогностні твердження лишаються умовними і потребують подальшої перевірки на підставі розширеного моніторингу та сценарного моделювання, і хоча карта спільна для цих видів, проте варто розуміти що тенденції щодо їх локалізації описані на початку підрозділу 3.1. все одно будуть зберігатися якщо взяти їх поодинці, і однозначно впливатимуть на динаміку їх поширення в майбутньому, і хоча вони поділяють багато вищеперечислених результатів в тій чи іншій мірі, це ще раз підкреслює те, що потрібні додаткові дослідження з іншими предикторами, якістю ґрунту, власне біом, рельєф однозначно тако відіграють свою роль в їх поширенні.

ВИСНОВКИ

Розроблено повний аналітичний цикл – від автоматизованого збору даних до статистичного моделювання фенологічних та просторових закономірностей. Поєднання підходів GAM/LMM і MaxEnt забезпечило комплексне розуміння взаємозв'язків між кліматичними параметрами та екологічною придатністю середовищ для *Phallus impudicus* і *Phallus hadriani*. Отримані результати мають наукове значення для мікологічного моніторингу, оцінки впливу кліматичних змін і розробки прогнозів поширення видів у майбутньому.

1. Здійснено збір та інтеграцію даних про поширення видів *Phallus impudicus* і *Phallus hadriani* на території України.

Використано записи з відкритих джерел громадянської науки (iNaturalist) та соціальної мережі Facebook (“Триби України”), отримані шляхом автоматизованих скриптів для пошуку та обробки контенту. Проведено стандартизацію дат, координат і назв таксонів. У результаті сформовано узгоджений набір присутностей із достовірними географічними координатами.

2. Проведено попередню обробку та валідацію спостережень. Застосовано фільтрацію за якістю, видалення дублікатів і просторове прорідження (spatial thinning) для усунення автокореляції та забезпечення репрезентативності. Додано кліматичні коваріати з баз WorldClim та ERA5 – 30-денні середні температури (T30), кумулятивні опади (P30), сумарні ефективні температури (degree-days) та висоту місцевості (elevation). Підготовлені дані стали основою для подальшого фенологічного й просторового аналізу.

3. Визначено фенологічні закономірності за допомогою узагальнених адитивних моделей (GAM) та лінійних змішаних моделей (LMM).

Аналіз показав істотний нелінійний вплив температури та опадів на дату появи плодових тіл (DOY). Для обох видів виявлено оптимальні температурні діапазони 15–20 °C, при яких частота спостережень максимальна, тоді як надлишкова сухість або перегрів спричиняють зсув фенологічних піків. Міжрічні

варіації, враховані у LMM, засвідчили статистично значущі кліматичні тренди, узгоджені з потеплінням останніх десятиліть.

4. Побудовано та проаналізовано просторові моделі розподілу (SDM) із використанням алгоритму MaxEnt.

Здійснено відбір фону (background) із урахуванням просторового зсуву вибірки (bias correction) та блокову крос-валідацію (block CV). Моделі виявили найбільшу ймовірність придатності серед лісових та узлісних ділянок із помірним зволоженням. Для *P. hadriani* характерніші південні й піщані локації, тоді як *P. impudicus* тяжіє до вологіших регіонів. Значення AUC > 0.85 підтверджують високу прогностичну якість моделей.

5. Виявлено просторово-часові зсуви фенології та кліматичну чутливість видів. Порівняльний аналіз часових рядів показав тенденцію до зсуву плодоношення в бік більш ранніх дат (у середньому на 5–8 днів за два десятиліття). Цей зсув корелює зі збільшенням середніх температур і зменшенням кількості опадів у вегетаційний період, що свідчить про адаптивну реакцію макроміцетів на зміну кліматичних умов.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Гриби України | Facebook [Електронний ресурс]. – URL : <https://www.facebook.com/groups/1054204501266549> (дата звернення: 03.10.2025).
2. Andrew C. et al. Congruency in fungal phenology patterns across dataset sources and scales //Fungal Ecology. – 2018. – Т. 32. – С. 9-17.
3. Andrew C. Not always optimal: Fungal fruiting triggers indicate climate sensitivity in cooler regions //Fungal Ecology. – 2025. – Т. 75. – С. 101416..
4. Andrew C. et al. Congruency in fungal phenology patterns across dataset sources and scales //Fungal Ecology. – 2018. – Т. 32. – С. 9-17.
5. Andrew C. et al. Explaining European fungal fruiting phenology with climate variability //Ecology. – 2018. – Т. 99. – №. 6. – С. 1306-1315.
6. Bidartondo M. I. et al. Climate change: Fungal responses and effects //In: Willis, Katherine J., editor. State of the World's Fungi 2018. Royal Botanical Gardens, Kew, UK. 62-69.. – 2018. – С. 62-69.
7. Buko V. et al. Polysaccharides of mushroom *phallus impudicus* mycelium: immunomodulating and wound healing properties // 现代食品科技. – 2019. – Т. 35. – №. 9. – С. 30–37.
8. Caldas L. A. et al. Immunomodulatory effect of polysaccharides from the mushroom-forming basidiomycete *Gymnopilus imperialis* (Agaricomycetes, Basidiomycota) //Pharmaceuticals. – 2022. – Т. 15. – №. 10. – С. 1179.
9. Cerrejón C. et al. Inconspicuous taxa in citizen science-based botanical research: actual contribution, limitations, and new opportunities for non-vascular cryptogams //Frontiers in Environmental Science. – 2025. – Т. 12. – С. 1448512..
10. *Clathrus ruber*, Red Cage or Lattice Fungus, identification // First-Nature.com [Електронний ресурс]. – URL : <https://www.first-nature.com/fungi/clathrus-ruber.php> (дата звернення: 19.10.2025).
11. Delmas C. E. L. et al. Monitoring the phenology of plant pathogenic fungi: why and how? //Biological Reviews. – 2024. – Т. 99. – №. 3. – С. 1075-1084.

12. Escobar D. F. E., Silveira F. A. O., Morellato L. P. C. Timing of seed dispersal and seed dormancy in Brazilian savanna: two solutions to face seasonality // *Annals of botany*. – 2018. – Т. 121. – №. 6. – С. 1197-1209.
13. *Geastrum fimbriatum*, Sessile Earthstar fungus // First-Nature.com [Электронный ресурс]. – URL : <https://www.first-nature.com/fungi/geastrum-fimbriatum.php> (дата звернення: 19.10.2025).
14. Haelewaters D. et al. The power of citizen science to advance fungal conservation // *Conservation Letters*. – 2024. – Т. 17. – №. 3. – С. e13013.
15. He H. et al. Inhibitory Effect and Mechanism of Hexanal on the Maturation of Peach-Shaped *Phallus impudicus* // *Journal of Fungi*. – 2025. – Т. 11. – №. 2. – С. 127.
16. Heluta V. P. Distribution of *Phallus hadriani* (Phallales, Basidiomycota) in Ukraine // *Ukrainian Botanical Journal*. – 2019. – Т. 76. – №. 1. – С. 24-30.
17. Kauserud H. et al. Warming-induced shift in European mushroom fruiting phenology // *Proceedings of the National Academy of Sciences*. – 2012. – Т. 109. – №. 36. – С. 14488-14493.
18. Kauserud H. et al. Climate change and spring-fruited fungi // *Proceedings of the Royal Society B: Biological Sciences*. – 2010. – Т. 277. – №. 1685. – С. 1169-1177.
19. Krah F. S. et al. Fungal fruit body assemblages are tougher in harsh microclimates // *Scientific reports*. – 2022. – Т. 12. – №. 1. – С. 1633.
20. Leal Filho W. et al. The environment as the first victim: The impacts of the war on the preservation areas in Ukraine // *Journal of Environmental Management*. – 2024. – Т. 364. – С. 121399.
21. Linchamps P. et al. Bioclimatic inference based on mammal community using machine learning regression models: perspectives for paleoecological studies // *Frontiers in Ecology and Evolution*. – 2023. – Т. 11. – С. 1178379.
22. More than 260 civilians killed after stepping on mines, explosives – Ukraine // Reuters. – (01.11.2023) [Электронный ресурс]. – URL : <https://www.reuters.com/world/europe/more-than-260-civilians-killed-after-stepping-mines-explosives-ukraine-2023-11-01/> (дата звернення: 19.10.2025).

23. *Mutinus caninus*, Dog Stinkhorn, identification // First-Nature.com [Электронный ресурс]. – URL : <https://www.first-nature.com/fungi/mutinus-caninus.php> (дата звернения: 19.10.2025).
24. *Phallus hadriani*, Dune Stinkhorn fungus // First-Nature.com [Электронный ресурс]. – URL : <https://www.first-nature.com/fungi/phallus-hadriani.php> (дата звернения: 19.10.2025).
25. *Phallus impudicus*, Stinkhorn fungus // First-Nature.com [Электронный ресурс]. — URL : <https://www.first-nature.com/fungi/phallus-impudicus.php> (дата звернения: 19.10.2025).
26. Pietras M., Kolanowska M., Selosse M. A. Quo vadis? Historical distribution and impact of climate change on the worldwide distribution of the Australasian fungus *Clathrus archeri* (Phallales, Basidiomycota) // *Mycological Progress*. – 2021. – Т. 20. – №. 3. – С. 299-311.
27. Piña-Páez C. et al. Diversity of gasteroid fungi in the Sierra de Mazatán, Sonora, Mexico // *The Southwestern Naturalist*. – 2013. – С. 351-356..
28. PP345/PP345: Stinkhorn Mushrooms (Agaricomycetes: Phallales: Phallaceae) // University of Florida IFAS Extension [Электронный ресурс]. — URL : <https://edis.ifas.ufl.edu/publication/PP345> (дата звернения: 19.10.2025).
29. Prylutskyi O., Kapets N. State-of-the-Art of iNaturalist as a Source of Data on Ukrainian Fungi // *Citizen Science: Theory and Practice*. – 2024. – Т. 9. – №. 1.
30. Pudil F., Uvira R., Janda V. Volatile compounds in stinkhorn (*Phallus impudicus* L. ex Pers.) at different stages of growth // *European Scientific Journal*. – 2014. – Т. 10. – №. 9. – С. 163-171.
31. Qazi A. W., Saqib Z., Zaman-ul-Haq M. Trends in species distribution modelling in context of rare and endemic plants: a systematic review // *Ecological Processes*. – 2022. – Т. 11. – №. 1. – С. 1-11.
32. Rosário I. T. et al. When do citizen scientists record biodiversity? Non-random temporal patterns of recording effort and associated factors // *People and Nature*. – 2025. – Т. 7. – №. 4. – С. 860-870.

33. Sakamoto Y. Influences of environmental factors on fruiting body induction, development and maturation in mushroom-forming fungi // *Fungal biology reviews*. – 2018. – Т. 32. – №. 4. – С. 236-248.
34. Staggemeier V. G. et al. The circular nature of recurrent life cycle events: a test comparing tropical and temperate phenology // *Journal of Ecology*. – 2020. – Т. 108. – №. 2. – С. 393-404.
35. The Editors of Encyclopaedia Britannica. Phenology // *Encyclopedia Britannica* (13.10.2019) [Электронный ресурс]. – URL : <https://www.britannica.com/science/phenology> (дата звернения: 19.10.2025).
36. Vogt-Schilb H. et al. Climate-induced long-term changes in the phenology of Mediterranean fungi // *Fungal Ecology*. – 2022. – Т. 60. – С. 101166.
37. Wu R. et al. Fungal spore seasons advanced across the US over two decades of climate change // *GeoHealth*. – 2025. – Т. 9. – №. 7. – С. e2024GH001323.
38. Yuan H. S., Wei Y. L., Wang X. G. Maxent modeling for predicting the potential distribution of Sanghuang, an important group of medicinal fungi in China // *Fungal Ecology*. – 2015. – Т. 17. – С. 140-145.
39. Zambonelli A. et al. Ectomycorrhizal fungal communities of edible ectomycorrhizal mushrooms // *Edible ectomycorrhizal mushrooms: Current knowledge and future prospects*. – Berlin, Heidelberg : Springer Berlin Heidelberg, 2013. – С. 105-124.

SUMMARY

Sivart D.D.

PHENOLOGY AND SPATIAL DISTRIBUTION OF *PHALLUS IMPUDICUS* AND *PHALLUS HADRIANI* IN UKRAINE BASED ON OPEN-SOURCE BIODIVERSITY DATA

This study analyzes the phenology and spatial distribution of *Phallus impudicus* and *Phallus hadriani* in Ukraine based on open observations and climate data. After filtering, 369 valid records were formed, to which climate predictors (T30, P30, degree days, elevation) were added.

The aim of the study was to assess the impact of weather conditions on fruiting times and to identify factors that shape the spatial suitability of the environment. For this purpose, circular statistics, generalized additive models (GAM), and range modeling (SDM/MaxEnt) were used.

It was found that both species bear fruit mainly in the second half of summer (average DOY ~214–216), and the key determinants of phenology are the 30-day average temperature and cumulative precipitation. Spatial models indicate an uneven response of species to climatic conditions and point to regions with the highest potential suitability. The novelty of the work lies in combining citizen science data with climatic predictors for quantitative analysis of these species in Ukraine.

Keywords: *Phallus impudicus*, *Phallus hadriani*, phenology, DOY, T30, P30, GAM, SDM, iNaturalist, Ukraine.