

Кандій Марія Олександрівна

Київський національний університет імені Тараса Шевченка, Київ, Україна,
e-mail: mariakandij@knu.ua, ORCID ID: 0009-0007-9358-4277

ВИКОРИСТАННЯ ДАНИХ ДИСТАНЦІЙНОГО ЗОНДУВАННЯ ЗЕМЛІ ДЛЯ ДОСЛІДЖЕННЯ ЩІЛЬНОСТІ НАДЗЕМНОЇ БІОМАСИ МАЛИХ УРБОГЕОСИСТЕМ (НА ПРИКЛАДІ МІСТА БУЧА)

Метою статті є розробка алгоритму модельного аналізу щільності надземної біомаси (англ. Above-Ground Biomass Density, AGBD) малих урбогеосистем на прикладі урбогеосистеми міста Буча на основі лідарних даних місії Global Ecosystem Dynamics Investigation (GEDI).

Методика. Моделювання проводилось із використанням методів машинного навчання, а саме методів Extreme Gradient Boosting та Random Forest. Алгоритм дослідження містив завантаження геопросторових даних, їхню обробку та проведення моделювання з оцінюванням точності результатів.

Результати. Було тестово застосовано, різні предиктори аналізу, серед них цифрову модель висот (ЦМВ) та спектральні індекси за даними супутнику Sentinel-2, та було протестовано два методи моделювання AGBD, а саме Extreme Gradient Boosting та Random Forest з урахуванням даних про системи землекористування та земельних покривів (LULC). За таких умов визначено, що за найкращий править метод моделювання Random Forest з використанням ЦМВ та спектральних індексів з геопросторовим розрізненням у 10 м. Було отримано дані про щільність надземної біомаси для малої урбогеосистеми м. Буча. Найбільша щільність надземної біомаси була зафіксована в зелених зонах з деревною рослинністю.

Наукова новизна. Аналіз попередніх досліджень вказує на відсутність загальноприйнятого переліку предикторів для моделювання щільності надземної біомаси. В дослідженні використано унікальний набір предикторів, визначений з огляду на особливості систем землекористування та земельних покривів малої урбогеосистеми. Інноваційним є використання шару-маски для урахування даних про системи землекористування та земельних покривів. Слід зазначити, що для території України оцінка щільності надземної біомаси на основі даних лідарних даних місії GEDI не проводилась, що також зумовлює наукову новизну дослідження.

Практична значимість. Використання даних лістанційного зондування Землі для оцінки щільності надземної біомаси дозволять замінити прямі вимірювання in situ, які є складними, а в деяких випадках і практично неможливими. Дані про щільність надземної біомаси є важливими для розуміння колообігу вуглецю, який в свою чергу впливає на кліматичну систему. На основі отриманих даних можна оцінити екосистемні урбогеосистеми задля досягнення кліматичної нейтральності.

Ключові слова: щільність надземної біомаси (AGBD), лідар-місія Global Ecosystem Dynamics Investigation (GEDI), метод Extreme Gradient Boosting, метод Random Forest, мала урбогеосистема, SRTM, Sentinel-2.

Kandii Mariia

Taras Shevchenko National University of Kyiv, Kyiv, Ukraine,
e-mail: mariakandij@knu.ua, <https://orcid.org/0009-0007-9358-4277>

ABOVE-GROUND BIOMASS DENSITY RESEARCH USING REMOTE SENSING DATA IN SMALL URBAN ECOSYSTEMS (ON THE EXAMPLE OF THE CITY OF BUCHA)

The purpose of this article is to develop an algorithm for modelling the above-ground biomass density (AGBD) of small urban geosystems using the example of the urban geosystem of the city of Bucha based on lidar data from the Global Ecosystem Dynamics Investigation (GEDI) mission.

Methodology. Modelling was performed using machine learning methods, namely Extreme Gradient Boosting and Random Forest. The research algorithm included loading geospatial data, processing it, and performing modelling with accuracy assessment of the results.

Results. Various analysis predictors were tested, including a digital elevation model (DEM) and spectral indices based on Sentinel-2 satellite data. Secondly, two AGBD modelling methods were tested, namely Extreme Gradient Boosting and Random Forest, taking into account data on land use and land cover (LULC) systems. Under these conditions, it was determined that the best modelling method is Random Forest using DEM and spectral indices with a geospatial resolution of 10 m. Data on above-ground biomass density for the small urban geosystem of Bucha were obtained. The highest above-ground biomass density was recorded in green areas with tree vegetation.

Scientific novelty. Analysis of previous studies indicates the absence of a generally accepted list of predictors for above-ground biomass density modelling. The study uses a unique set of predictors, determined taking into account the characteristics of a small urban geosystem LULC. The use of a layer mask to take into account data on LULC is an innovative approach. It should be noted that no assessment of above-ground biomass

density based on GEDI lidar data has been conducted for the territory of Ukraine, which also contributes to the scientific novelty of the study.

Practical significance. The use of Earth observation data to estimate above-ground biomass density enable the replacement of direct in situ measurements, which are difficult and, in some cases, practically impossible. Data on above-ground biomass density are important for understanding the carbon cycle, which in turn affects the climate system. Based on the data obtained, ecosystem services to achieve climate neutrality in urban geosystems can be assessed.

Keywords: above-ground biomass density (AGBD), Global Ecosystem Dynamics Investigation (GEDI) lidar mission, Extreme Gradient Boosting method, Random Forest method, small urban geosystem, SRTM, Sentinel-2.

Постановка проблеми. Глобальна зміна клімату на сьогоднішній день залишається найсерйознішим викликом для життєдіяльності людства. Одним із чинників, які істотно впливають на кліматичну нейтральність, є глобальний колообіг вуглецю. Для його стабілізації необхідним є досягнення нетто-нульових викидів, тобто такого стану, за якого обсяги викидів парникових газів цілком компенсуються обсягами їхнього видалення з атмосфери. З огляду на це, важливим є дослідження процесів секвестрації та депонування вуглецю в різних геосистемах, які можна оцінити на основі даних про щільність надземної біомаси. Оцінювання такої щільності за допомогою прямих вимірювань *in situ* є дуже складним, а досить часто і практично неможливим чи недоцільним завданням. Саме тому вельми перспективним для дослідження надземної біомаси є використання даних дистанційного зондування Землі (ДЗЗ). Крім того, такі дослідження є актуальними щодо урбогеосистем, зокрема малих, через їхню високу міру антропоїзації, яка перетворює ці геосистеми на істотне джерело необхідних для видалення парникових газів.

Аналіз останніх досліджень і публікацій. У сучасній науковій практиці згідно з (Wang J. et al., 2025) зазвичай виділяють три типи даних, які використовуються при непрямих методах оцінки надземної біомаси, а саме дані, отримані мікрохвильовим радаром, лідаром (LiDAR) та пасивними сенсорами як оптичні зображення. У нашому дослідженні було вирішено слухним використанням даних лідар-місії Global Ecosystem Dynamics Investigation (далі – GEDI), тому досвід використання інших джерел даних не розглядався.

У сучасній науковій літературі (Wang J. et al., 2025, Getachew et al., 2025, Lutz et al., 2024, Zhao et al., 2024, Wang Y. et al., 2024, Liang et al., 2023) широко використовуються методи машинного навчання для побудови континуальної поверхні на основі точкових даних про надземну біомасу. Сутність даних методів полягає у передбаченні щільності надземної біомаси (Above-Ground Biomass Density, далі, аббревіатурно, AGBD) на тих ділянках, для яких відсутні дані, на основі використання континуальних даних, які впливають на значення AGBD. Таким чином, лідарні дані виступають як залежна змінна, а континуальні дані – як незалежні змінні (предиктори) (Wang J. et al., 2025). У наукових першоджерелах відсутній загальноприйнятий перелік таких предикторів, утім, різними дослідниками зазвичай використовуються дві групи незалежних змінних, а саме параметри рельєфу та дані оптичного знімання, в основному спектральні індекси (табл.1).

Таблиця 1.

Використання предикторів AGBD в дослідженнях надземної біомаси

| автори, рік | параметри рельєфу | | спектральні індекси | |
|---|-------------------|---|---|---|
| | джерело даних | параметри | джерело даних | індекси |
| Birhan Getachew Tikuye, Ram Lakhan Ray, 2025 | SRTM | - цифрова модель висот - експозиція схилів | Landsat-9 Operational Land Imager (OLI) | NDVI, EVI, PVI, NBR, NDWI, NDBI, MSI |
| Jiarun Wang, Chengzhi Xiang and Ailin Liang, 2025 | SRTM | - цифрова модель висот - похил поверхні - експозиція схилів | Sentinel-2 | Усі можливі комбінації різниці каналів Sentinel-2 попарно (45 індексів) |
| Natasha Lutz, Pedro Rodriguez-Veiga and Imma Oliveras Menor, 2024 | SRTM | - цифрова модель висот - похил поверхні - експозиція схилів | Sentinel 2, Landsat 8 | NDVI, GNDVI, NDVI-RE, MSAVI, EVI, NDWI |

| | | | | |
|--|------|---|-----------------------|--|
| Xuedi Zhao, Wenmin Hu, Jiang Han, Wei Wei, Jiaying Xu, 2024 | SRTM | - цифрова модель висот | Sentinel-2, Landsat 8 | NDVI, GNDVI, EVI, NBR, RVI, DVI, ARVI |
| Yingchen Wang, Hongtao Wang, Cheng Wang, Shuting Zhang, Rongxi Wang, Shaohui Wang, Jingjing Duan, 2024 | SRTM | - цифрова модель висот - похил поверхні - експозиція схилів | Sentinel-2 | NDVI, RVI, EVI, DVI, MNDWI, RGVI, SAVI, SVI |
| Mengyu Liang, Laura Duncanson, Julie A. Silva, Fernando Sedano, 2023 | SRTM | - цифрова модель висот - похил поверхні | Landsat 5,8 | NDVI, NBR, NDMI, EVI, SAVI, ITZ, TC brightness, TC angle, TC Disturbance |

Примітка. Повні назви параметрів, аббревіатури чи скорочення яких використано у таблиці: SRTM – Shuttle Radar Topography Mission; NDVI – Normalized Difference Vegetation Index; EVI – Enhanced Vegetation Index, PVI – Perpendicular Vegetation Index, NBR – Normalized Burn Ratio, NDWI – Normalized Difference Water Index, NDBI – Normalised Difference Built-up Index, MSI – Moisture Stress Index, GNDVI – Green Normalised Differenced Vegetation Index, NDVI-RE – Normalized Difference Vegetation Index Red-edge, MSAVI – Modified Soil Adjusted Vegetation Index, RVI – Radar Vegetation Index, DVI – Difference Vegetation Index, ARVI – Atmospherically Resistant Vegetation Index, MNDWI – Modified Normalized Difference Water Index, RGVI – Red-Green Vegetation Index, SAVI – Soil-Adjusted Vegetation Index, SVI – Shadow Vegetation Index, NDMI – Normalized Difference Moisture Index, IFZ – Integrated Forest Zscore, TC brightness – Tasseled Cap Brightness, TC angle – Tasseled Cap angle, TC Disturbance – Tasseled Cap Disturbance (див. детальніші пояснення далі за текстом).

Зміст таблиці 1 дозволяє зробити такі узагальнення:

1. У всіх розглянутих дослідженнях використовуються дані проєкту Shuttle Radar Topography Mission (далі SRTM). Набір параметрів при цьому є майже тотожним: усі праці враховують цифрову модель висот, а також одну або дві похідні від неї: похил поверхні та/або експозицію схилів. Обґрунтовуючи доцільність використання цифрової моделі висот (далі – ЦМВ) для передбачення AGBD, дослідники наголошують передусім на впливі параметрів рельєфу на умови розвитку надземної рослинності, насамперед деревної.

2. Частота використання даних супутників Landsat та Sentinel є приблизно однаковою.

3. З вегетаційних індексів усі дослідники використовують NDVI, переважна більшість застосовує EVI, подальший же набір індексів відрізняється. Отже, не існує стандартних та загальноприйнятих вимог щодо використання чітко визначеного переліку індексів, які розглядаються. Це дозволяє обирати їх з огляду на об'єкт та мету дослідження.

За геопросторові області дослідження AGBD на основі лідарних даних правили одиниці адміністративно-територіального та природничого поділу різних рівнів, а саме:

1) держави: КНР (*Wang J., Xiang C. and Liang A., 2025*);

2) адміністративно-територіального поділу держав: округ Сонома, США (*Wang Y et al., 2024*), міський округ Сюйчжоу, КНР (*Zhao X. et al., 2024*);

3) біогеографічного районування: два регіони та один субрегіон біогеографічного районування Австралії (*Lutz N. et al., 2024*);

4) річкових басейнів та їхніх частин: басейн нижньої Бразос, США (*Tikuye B., Ray R., 2025*);

5) невеликих територій в межах населених пунктів: кампус Університету Британської Колумбії, Канада (*Gülçin D., Bosch C., 2021*), кампус Натан університету Гріффіт, Австралія (*Thinley J., Ndehedehe Ch., 2025*);

6) міських зелених зон: 2 заліснених масиви в провінції Юньнань (*Wang C. et al., 2024*).

Виділення невирішених раніше частин загальної проблеми. По-перше, серед дослідників відсутній уніфікований набір предикторів щільності надземної біомаси. По-друге, дослідження проводились здебільшого для територій з щільним рослинним покривом. Малі урбогеосистеми не виступали об'єктами дослідження, в той час як особливості їхньої просторової організації потребують розробки специфічних алгоритмів дослідження. По-третє, для території України у журналах, проіндексованих базою Web of Science, відсутні публікації, присвячені оцінці AGBD на основі даних GEDI. У журналах, проіндексованих базою Scopus, наявні приклади використання даних GEDI для дослідження структурних характеристик пралісів в Українських Карпатах (*Spracklen B. and Spracklen D.V., 2021*) та прогнозування лісових пожеж (*Myroniuk V. et al., 2023*). Оцінки ж щільності надземної біомаси на основі даних GEDI для України не проводились, а отже існує необхідність дослідження особливостей такого застосування цих даних для території нашої країни.

Формулювання цілей статті. Ціллю дослідження є розробка алгоритму визначення щільності AGBD малих урбогеосистем на основі даних місії GEDI. Для досягнення поставленої мети необхідним є виконання наступних задач, а саме: (1) визначення предикторів, (2) обрання моделі, (3) оцінка точності моделі, (4) у разі виявлення значних неточностей – введення додаткових предикторів. За об'єкт дослідження правлять малі урбогеосистеми на прикладі урбогеосистеми міста Буча Київської області України. Предметом дослідження є щільність надземної біомаси.

Вклад основного матеріалу. Передусім розглянемо використані дані.

Зрозуміло, що найбільш достовірним джерелом даних про надземну біомасу залишаються прямі польові вимірювання. Утім, такі дослідження є дуже витратними за часом, людськими ресурсами і, наслідково, вартістю. З огляду на це, проведення зазначених розвідок є доцільним лише на невеликих дослідних ділянках (Tikuye B., Ray R., 2025, Kandii M., Samoilenko, V., 2025). З розвитком технологій ДЗЗ з'явилась можливість отримання із певною періодичністю параметрів рослинного покриву урбогеосистем, різноманітних як за площею, так і за кількістю для значних за розмірами територій.

Технології ДЗЗ використовують пасивні та активні сенсори. Перші дозволяють отримати дані лише про верхній шар рослинності, тому для оцінки AGBD є малоприматними і потребують додаткових вимірів, переважно наземних. Активні сенсори відстежують вертикальну структуру рослинного покриву, і тому можуть бути використані для оцінки AGBD.

Широкого застосування для оцінки надземної біомаси набули дані місії Global Ecosystem Dynamics Investigation (GEDI) – багатопроменевого лідара Міжнародної космічної станції. Інструмент складається з трьох лазерів, які продукують 8 променів трансект. Це дозволяє отримувати одночасно 8 точкових даних розміром приблизно 25x25 м, розташованих приблизно через кожні 60 м уздовж трансекти при відстані між трансектами у межах 600 м (Dubayah, R.O. et al., 2022, ORNL DAAC, 2025). Таким чином, ці дані є дискретними. Початково отримуються тривимірні вертикальні параметри рослинного покриву, на основі яких за допомогою відповідних алгоритмів визначаються прогнозована щільність надземної біомаси (у Мг/га, тобто у тонах на гектар) та інші допоміжні параметри, зокрема стандартну похибку. Дані GEDI підпадають під обробку чотирьох рівнів. У нашому модельному аналізі було використано дані рівня 4A, який містить прогнозовану щільність надземної біомаси на точковому рівні (Footprint Level Aboveground Biomass Density). Такі відомості доступні для завантаження у HDF5 (*.h5) форматі.

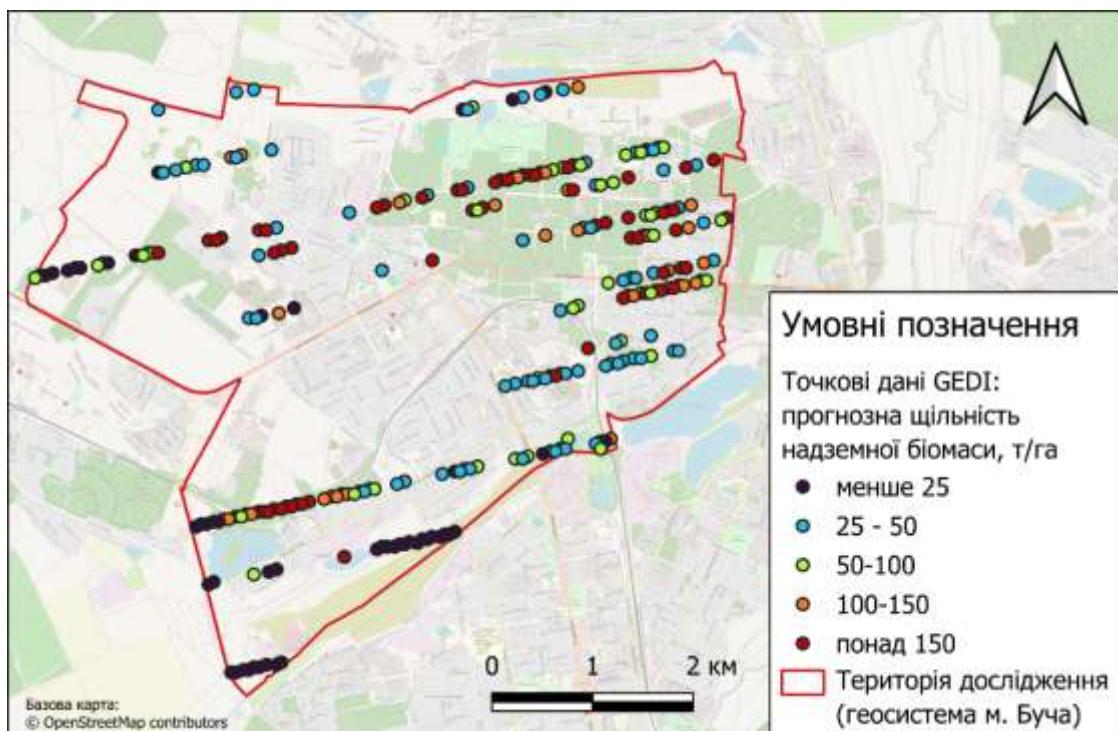


Рис. 1. Розташування точок із прогнозованою щільністю надземної біомаси урбогеосистеми Бучі за вихідними даними GEDI (за NASA's Earth science data, 2025)

Пошук даних GEDI було здійснено на платформі Earthdata Search (*NASA's Earth science data*, 2025). Оскільки для території дослідження характерні коливання щільності надземної біомаси в залежності від сезону, пошук здійснювався за останніми наявними даними в період максимальної вегетації, а саме з 15 травня по 15 вересня. Дані такого часового проміжку були наявні за 2024 р. у кількості 6 файлів. Такі дані було завантажено, після чого в середовищі Google Colaboratory (Google Colab) їх було обрізано за межами досліджуваної малої урбогеосистеми Бучі та переведено у формат GeoTiff. За таких умов було отримано дані в 263 точках (footprints). На рис. 1 показано розташування точок та прогнозна щільність біомаси. Цифровий шар рисунку було векторизовано для зручності візуалізації, для всіх же подальших обчислень використовувався растровий формат геоданих.

Як дані оптичних сенсорів було використано супутникові дані Sentinel-2 рівня L2A, тобто дані з атмосферною корекцією, які відображають відбивальну здатність поверхні, а не відбивальну здатність на межі атмосфери (*Copernicus Sentinel-2*, 2021). Даним Sentinel була надана перевага над даними Landsat з огляду на більше геопросторове розрізнення, що є важливим при дослідженні урбогеосистем. Пошук даних здійснювався за допомогою браузеру Copernicus (*Copernicus Data Space Ecosystem Browser*, 2025) для часового проміжку, аналогічного отриманим даним GEDI. Серед наявних знімків Sentinel для території дослідження було обрано один із найменшим покриттям хмарами, а саме космоснімок за 17.08.2024 р.

За цифрову модель висот правили дані проєкту Shuttle Radar Topography Mission (SRTM), завантажені на платформі Earth Explorer (*United States Geological Survey*, 2025).

Дані про системи землекористування та земельних покривів (LULC) було отримано з генерального плану м. Бучі, завантажений з офіційного сайту Бучанської міської ради (Шпилевський та ін, 2015). Такі геодані було оцифровано ГІС-інструментарієм QGIS та актуалізовано із використанням композитних зображень Sentinel-2 та інших джерел інформації (*дані Google Satellite*, 2025, *Open Street Map*, 2025).

Методи. Для проведення запланованих досліджень було розроблено їхню схему (рис.2). Вона передбачала три етапи: отримання геопросторових даних, обробка цих даних та моделювання щільності надземної біомаси. Етап отримання геоданих вже описано при характеристиці використаних даних. На етапі ж обробки ці дані було підготовлено до моделювання, а саме: точкові дані GEDI переведено у формат GeoTiff; обчислено спектральні індекси, похил поверхні та експозицію схилів; дані про системи LULC переведено у растровий формат та зроблено шар-маску (механізм чого описано далі у результатах дослідження).

На етапі моделювання для інтерполяції точкових даних про AGBD було здійснено регресійне моделювання з використанням моделей обраних методів машинного навчання, а саме Extreme Gradient Boosting (XGBoost) та Random Forest (RF). Обробка даних GEDI, застосування моделей методів машинного навчання та оцінка їхньої точності за допомогою відповідних статистичних параметрів були здійснені шляхом написання відповідних кодів мовою Python. За середовище застосування моделей правило хмарне середовище Google Colaboratory (скорочено - Google Colab) (*Google Colaboratory*, 2025). Доцільність використання зазначеного хмарного середовища зумовлена головним чином тим, що обробка даних, використаних в дослідженнях, потребує значних ресурсів.

Оцінка точності моделей, які тестувалися, тобто відхилень у лінійних залежностях між прогнозованими та зафіксованими результатами, проводилась шляхом обчислення середньоквадратичної помилки (Root Mean Square Error, RMSE) та коефіцієнта детермінації (R^2) зазначених залежностей. При цьому враховувалося, що кращою є, зрозуміло, модель з нижчою RMSE, а значення коефіцієнта детермінації маркують регресійний зв'язок між прогнозованими та зафіксованими параметрами як задовільний при $0,36 < R^2 \leq 0,64$ і добрий при $R^2 > 0,64$ (*Самойленко, Діброва*, 2025).

Достовірність результатів моделювання щільності біомаси також перевірялася і коригувалася з огляду на те, до яких систем LULC були приурочені ці результати.

Візуалізація геопросторових даних проводилась ГІС-інструментарієм QGIS.

Правильний добір предикторів є визначальним для отримання найбільш достовірних результатів, і повинен базуватись на характеристиках досліджуваної території та меті дослідження. Крім того, за об'єкт дослідження правили малі урбогеосистеми, які вирізняються певними особливостями систем землекористування та земельних покривів, які впливають на AGBD.

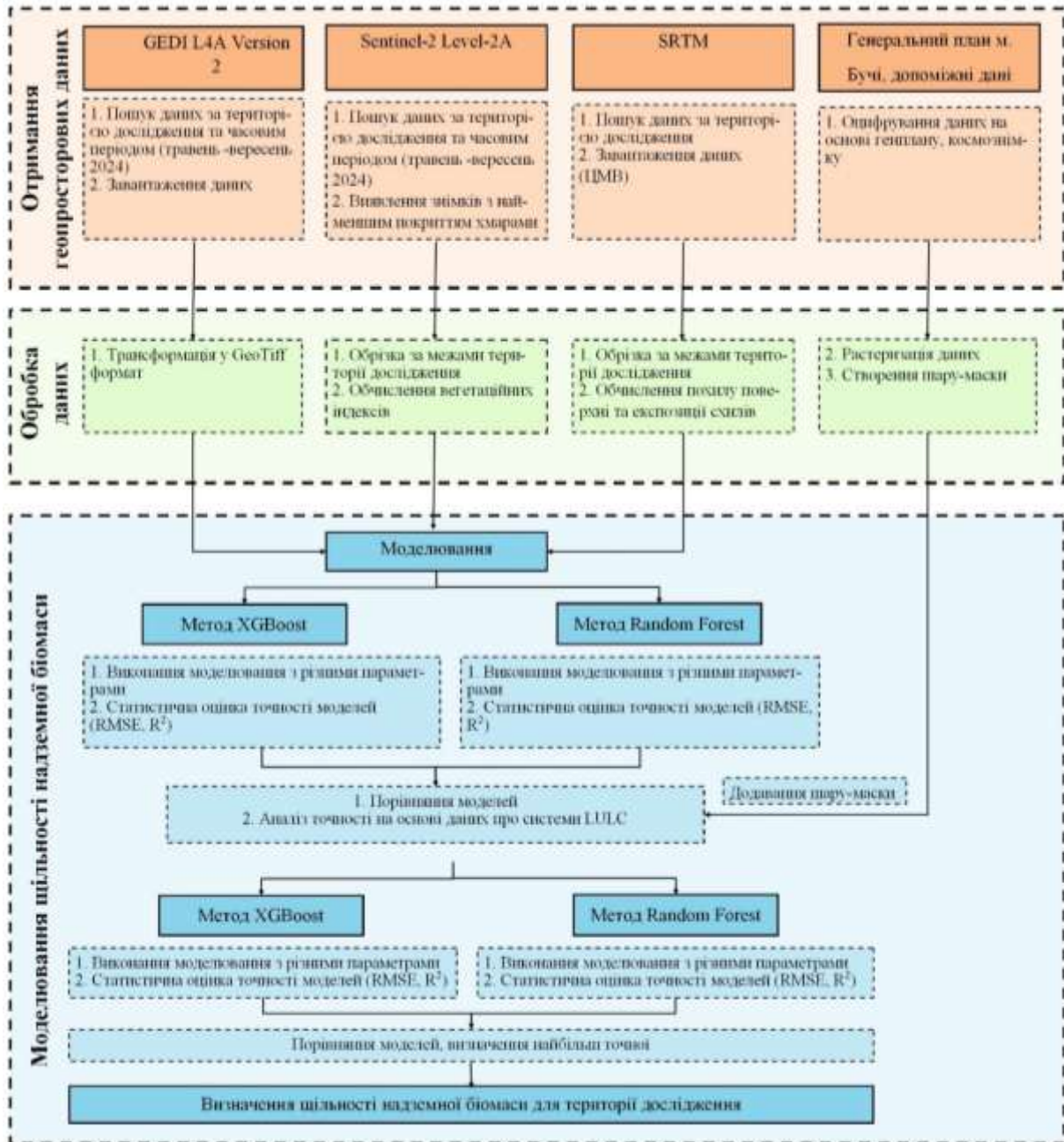


Рис. 2. Схема досліджень

Так, по-перше, надземну біомасу зазначених урбогеосистем зосереджено переважно в деревній рослинності. Деревна рослинність цих урбогеосистем зосереджено у зелених зонах загального користування (парки, сквери, міські ліси тощо), територіях обмеженого користування (зелені насадження на території певних установ та озеленені ділянки в межах забудови), а також територіях спеціального призначення (охоронні зони різних об'єктів). До того ж щільність надземної біомаси залежить від вертикальної структури та стану рослинності. Відповідно, регресійний аналіз повинен містити предиктори, які описують таку структуру та стан.

По-друге, на території малих урбогеосистем зазвичай розташовано штучні та природні водні об'єкти. Щільність надземної біомаси для них практично дорівнює нулю, що і має бути враховано через відповідний предиктор.

По-третє, малі урбогеосистеми вирізняються наявністю та високою щільністю забудови. З огляду на невелике поширення в Україні вертикального та дахового озеленення, щільність надземної біомаси для комірок растра, які приурочено до будівель, близька до нуля. У той же час, на ділянках відкритого ґрунту, які за своїми спектральними характеристиками є близькими до забудови, наявність рослинності або її поява є можливою, що призводить до значень надземної

біомаси, відмінних від нуля. Це спричинює необхідність запровадження предиктора, який допоможе відрізнити зазначені категорії земельних покриттів.

Зважаючи на вищезазначене, як предиктори було обрано такі спектральні індекси, як:

1. NDVI (Normalized Difference Vegetation Index) – нормалізований диференційний вегетаційний індекс або стандартизований індекс відмінностей рослинного покриття, який визначає стан останнього, зокрема його щільність. Є одним із найпоширеніших з т.зв. вегетаційних індексів. Формула для обчислення NDVI на основі даних Sentinel 2 має вигляд:

$$NDVI = \frac{B8 - B4}{B8 + B4}$$

де тут і далі: Bx – відповідний номер каналу Sentinel 2 L2A.

2. EVI (Enhanced Vegetation Index) – покращений вегетаційний індекс. Він також відображає стан рослинності, але завдяки додаванню синього каналу менше залежить від параметрів атмосфери та ґрунту і є більш чутливим для густої рослинності. Цей індекс є важливим при оцінці AGBD на лісовкритих територіях. Хоча урбогеосистеми зазвичай не вирізняються густим рослинним покритвом, використання EVI є доцільним для покращення моделювання AGBD систем LULC з близькою до природної рослинністю (міські ліси, лісопарки, об'єкти природно-заповідного фонду). EVI обчислюється за формулою:

$$EVI = 2.5 * \left(\frac{B8 - B4}{B8 + 6 * B4 - 7.5 * B2 + 1} \right)$$

3. SAVI (Soil-Adjusted Vegetation Index) – пристосований до ґрунту вегетаційний індекс, який допомагає знизити вплив відбивання опромінюваного ґрунту на спектральні значення NDVI. Цей індекс зазвичай використовується для дослідження систем LULC із розрідженим рослинним покритвом. Формула для обчислення SAVI записується як:

$$SAVI = \frac{(B8 - B4) * (1 + L)}{(B8 + B4 - L)}$$

де L – показник, який враховує щільність рослинного покриття і може набирати значення від 0 (щільна рослинність) до 1 (дуже розріджена рослинність). Досить часто оперують середнім значенням цього індексу (0,5).

4. NDWI (Normalized Difference Water Index) – нормалізований диференційний водний індекс, який використовується для ідентифікації поверхневих водних об'єктів. Формула для обчислення NDWI має вигляд:

$$NDWI = \frac{B3 - B8}{B3 + B8}$$

5. NDBI (Normalized Difference Built-up Index) – нормалізований диференційний індекс забудови, який використовується для ідентифікації штучних поверхонь, таких як забудова, дороги тощо. При дослідженні урбогеосистем він є важливим, оскільки дозволяє відрізнити водонепроникні поверхні від ґрунту. Формула для обчислення NDBI виглядає як:

$$NDBI = \frac{B11 - B8}{B11 + B8}$$

Як вже зазначалося, аналіз попередніх досліджень (Wang J. et al., 2025, Getachew et al., 2025, Lutz et al., 2024, Zhao et al., 2024, Wang Y. et al., 2024, Liang et al., 2023) показав, що для передбачення AGBD використовуються не тільки спектральні індекси, а й ЦМВ. Обґрунтовуючи використання ЦМВ для передбачення AGBD, дослідники наголошують на впливі параметрів рельєфу на умови росту дерев. Утім, при дослідженні малих урбогеосистем дані SRTM щодо поверхні їхнього рельєфу можуть мати т.зв. параметри, що заважають. Мова йде про випадки, коли замість відображення власне рельєфу, тобто висот поверхні Землі, дані SRTM будуть хибно відображати висоти певних об'єктів на цій поверхні, а саме будинків, мостів, зелених насаджень тощо, що може спотворити модельні результати. За таких застережень було проведено два альтернативних варіанти моделювання біомаси: з використанням ЦМВ за даними SRTM, а саме щодо висоти поверхні, ухилу та експозиції, та без їх використання.

Таким чином, для кожного обраного набору предикторів було здійснено моделювання щільності надземної біомаси малої урбогеосистеми м. Буча обох методів, тобто XGBoost та RF. Крім того, моделювання виконувалося і для різного геопросторового розрізнення (ГР) вихідних даних, а саме для ГР=10 м і ГР=30 м.

Статистичні показники точності проведеного моделювання (табл. 2) засвідчили таке. Найнижчі помилки моделювання отримуються при використанні як предикторів ЦМВ та вегетаційних індексів при ГР у 10 м (RMSE=47,754 для XGBoost та 53,788 для RF). При ГР у 30 м більш точні результати дає моделювання без використання ЦМВ, при ГР у 10 м – моделювання

з використанням ЦМВ. За однакових предикторів та однакового ГР більш точні результати дає застосування моделей методу RF. Окрім статистичних параметрів, про достовірність результатів моделювання може свідчити область значень цих результатів. Так, використання методу XGBoost дає від'ємні значення AGBD, що є неможливим. Натомість метод RF дає тільки додатні значення результатів, що є вагомим аргументом на користь його застосування. З огляду на все вищезазначене, найкращими за отриманими результатами виявилось моделювання AGBD методом RF з використанням ЦМВ та спектральних індексів з ГР у 10 м.

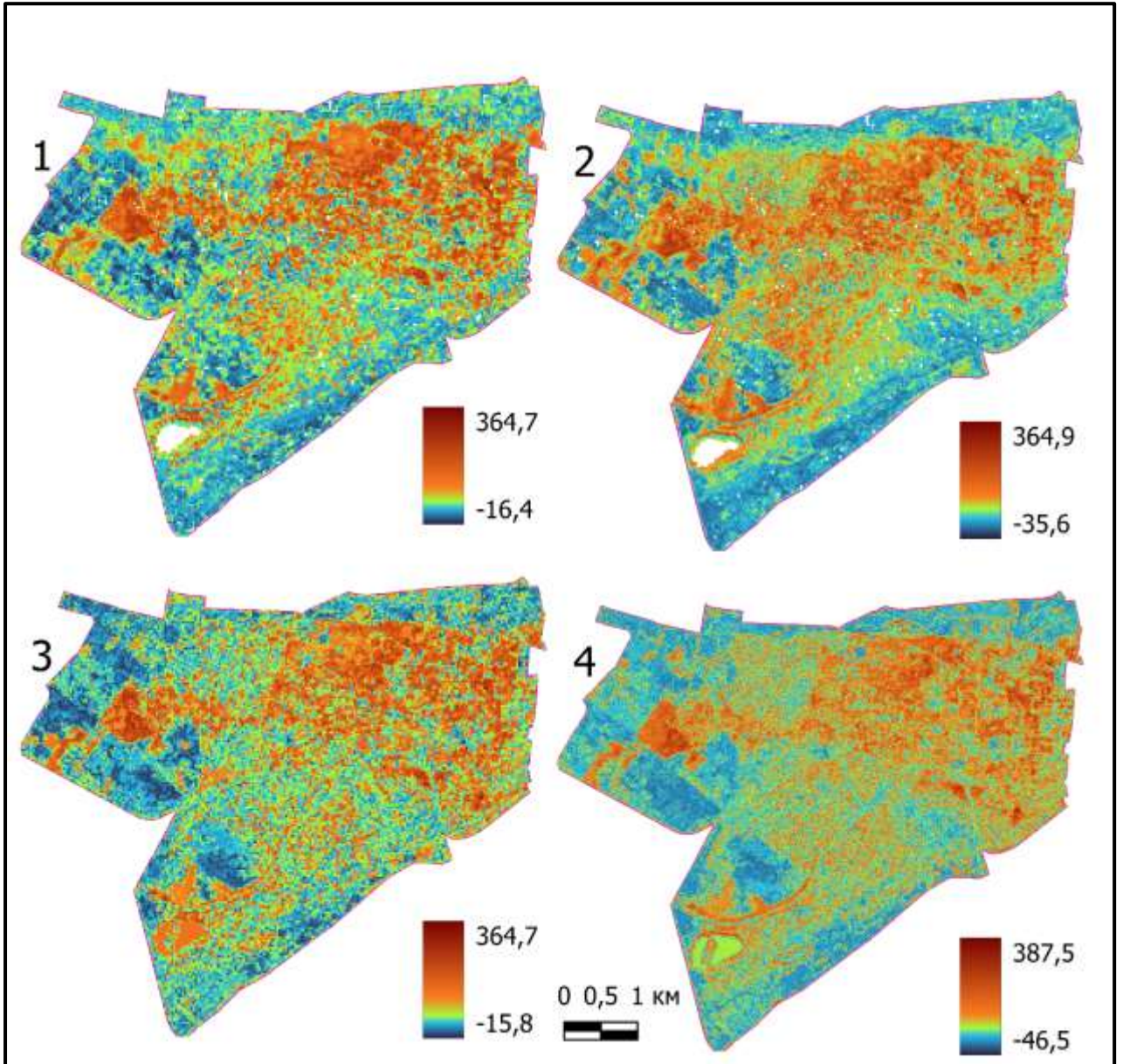


Рис. 3. Моделювання щільності надземної біомаси (AGBD) урбогеосистеми м. Буча за допомогою методу XGBoost: 1 – з використанням спектральних індексів та ЦМВ, ГР=30 м; 2 – з використанням спектральних індексів та ЦМВ, ГР=10 м; 3 – з використанням спектральних індексів, ГР=30 м; 4 – з використанням спектральних індексів, ГР=10 м.

Візуалізація та аналіз цифрових карт із змодельованою для геосистеми Бучі AGBD (рис.3-4) дозволяє відзначити такі закономірності. При використанні як предикторів ЦМВ та спектральних індексів для водойми на південному заході геосистеми значення AGBD обчислюються як "відсутні дані" (Рис. 3.1, 3.2 та Рис 4.1, 4.2). При використанні тільки спектральних індексів для зазначеної водойми величини AGBD набирають додатних значень, що є хибним результатом (Рис. 3.3, 3.4 та Рис 4.3, 4.4). Застосування предикторів ЦМВ та спектральних індексів призводить до появи окремих комірок растру із відсутніми даними досить

рівномірно по всій території, що утім не є критичним. Використання лише спектральних індексів до появи таких комірок не призводить. Метод RF у порівнянні з оцінками за XGBoost дає дещо більші показники AGBD для ділянок із щільною здоровою рослинністю і дещо менші для ділянок з розрідженою рослинністю. У цілому, обидва методи дають схожий геопросторовий розподіл змодельованих значень AGBD (Рис. 3, 4).

Таблиця 2.

Показники точності моделювання щільності надземної біомаси (AGBD) урбогеосистеми м. Буча для різних предикторів, модельних методів і геопросторового розрізнення

| Предиктори | Моделі методу | ГР | RMSE, т/га | R ² | Область значень, т/га |
|--|---------------|-------------|---------------|----------------|-----------------------|
| SRTM, Slope, Aspect, NDVI, EVI, NDWI, NDBI, SAVI | XGBoost | 30 м | 68,808 | 0,488 | -16,4...364,7 |
| | | 10 м | 47,754 | 0,707 | -35,6...364,9 |
| | Random Forest | 30 м | 67,153 | 0,513 | 2,8...307,1 |
| | | 10 м | 53,788 | 0,629 | 1,3...344,4 |
| NDVI, EVI, NDWI, NDBI, SAVI | XGBoost | 30 м | 63,118 | 0,489 | -15,7...364,7 |
| | | 10 м | 68,358 | 0,375 | -46,5...387,5 |
| | Random Forest | 30 м | 58,756 | 0,558 | 1,8...303,8 |
| | | 10 м | 67,699 | 0,387 | 1,5...341,7 |

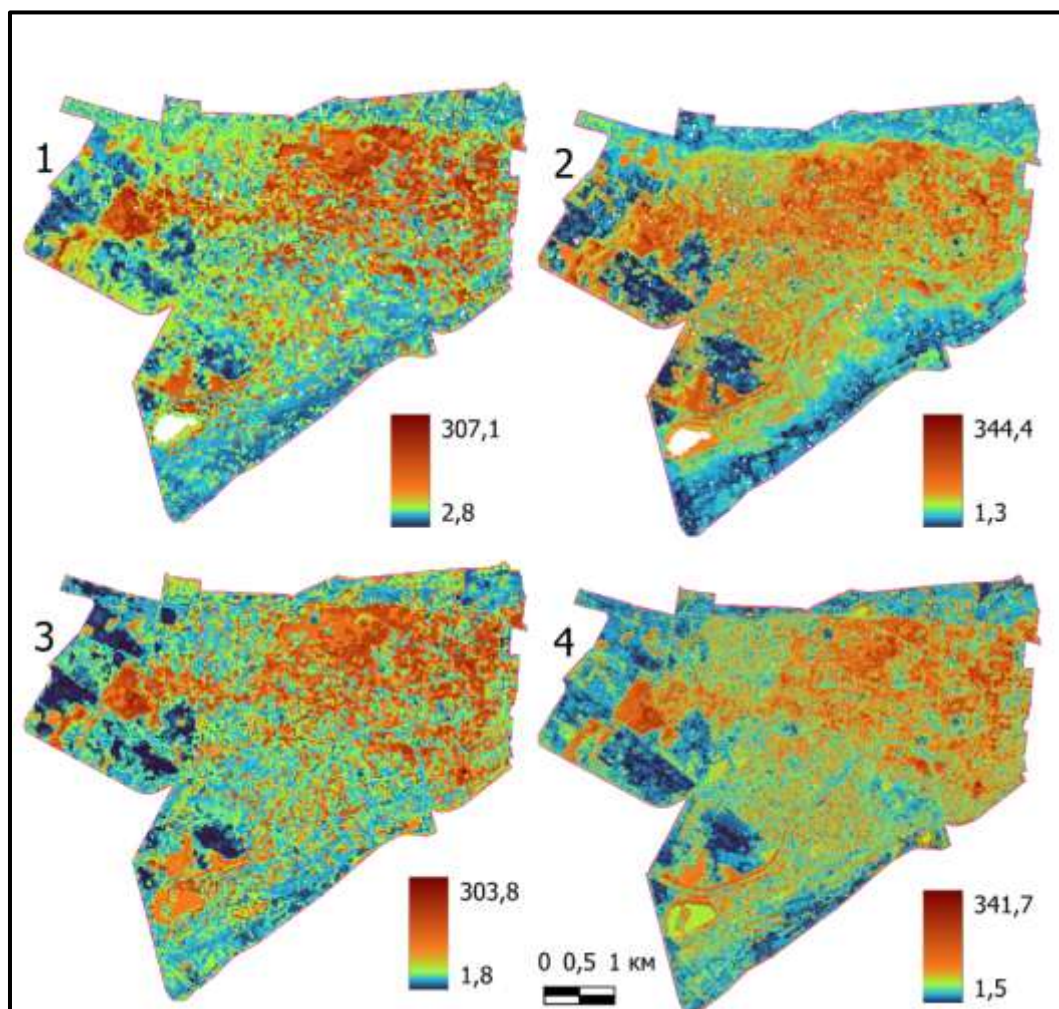


Рис. 4. Моделювання щільності надземної біомаси (AGBD) урбогеосистеми м. Буча за допомогою методу RF: 1 – з використанням спектральних індексів та ЦМВ, ГР=30 м; 2 – з використанням спектральних індексів та ЦМВ, ГР=10 м; 3 – з використанням спектральних індексів, ГР=30 м; 4 – з використанням спектральних індексів, ГР=10 м

Пошарове зіставлення отриманих результатів моделювання щільності надземної біомаси із даними про системи землекористування та земельних покривів показало, що більшість моделей призвели до появи додатних значень AGBD для водних об'єктів та автодоріг і залізничних колій, для яких AGBD повинна бути близькою до нуля. При цьому такі оціночні значення виявились подекуди навіть більшими за значення для трав'яного покриву, що, зрозуміло також не відповідає дійсності.

Для розв'язання щойно зазначеної проблеми було створено растровий шар-маску з геопросторовим розрізненням 10 м та 30 м. У цьому шарі комірка растра, які відповідають водним об'єктам та дорогам, було присвоєно значення "0", а решті комірок – значення "1". Такий шар-маску було інтегровано в моделі з умовою, що для комірок із значенням "0" значення AGBD становить "0", а для комірок із значенням "1" значення AGBD моделюються за відповідним алгоритмом.

За таких умов показники точності моделювання (Табл. 3) дещо змінились лише для методу XGBoost з використанням ЦМВ. Найкращі ж результати, як і за табл.2, дає моделювання AGBD методом RF з використанням ЦМВ та спектральних індексів з ГР у 10 м.

Табл. 3.

Показники точності моделювання щільності надземної біомаси (AGBD) урбогеосистеми м. Буча для різних предикторів, модельних методів і геопросторового розрізнення із застосуванням шару-маски

| Предиктори | Моделі методу | ГР | RMSE, т/га | R ² | Область значень, т/га |
|---|---------------|-------------|---------------|----------------|-----------------------|
| SRTM, Slope, Aspect, NDVI, EVI, NDWI, NDBI, SAVI, шар-маска | XGBoost | 30 м | 65,189 | 0,541 | -15,7...364,7 |
| | | 10 м | 48,664 | 0,696 | -40,5...364,6 |
| | Random Forest | 30 м | 67,153 | 0,513 | 0...307,1 |
| | | 10 м | 53,788 | 0,629 | 0...344,4 |
| NDVI, EVI, NDWI, NDBI, SAVI, шар-маска | XGBoost | 30 м | 63,118 | 0,489 | -15,7...364,7 |
| | | 10 м | 68,358 | 0,375 | -46,5...387,5 |
| | Random Forest | 30 м | 58,756 | 0,558 | 0...303,8 |
| | | 10 м | 67,699 | 0,387 | 0...341,7 |

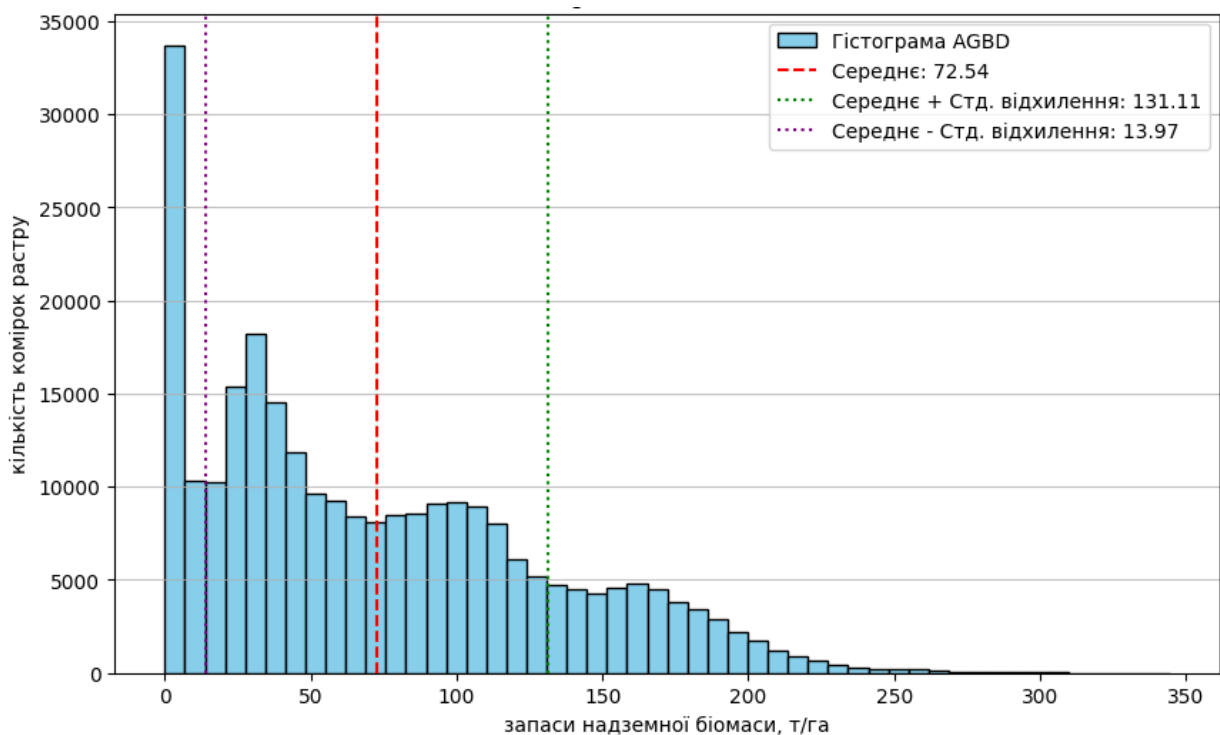


Рис. 5. Гістограма розподілу за комітками растра значень AGBD, змодельованих для урбогеосистеми м. Буча

Порівняння наведених в табл.2-3 статистичних параметрів точності з такими ж параметрами моделей інших дослідників дозволяє зробити висновок про діапазони отриманих похибок. Зокрема, у праці (Wang Y. et al., 2024) значення RMSE для різних моделей становили від 49,56 до 78,05 т/га, в праці (Wang J. et al., 2025) – від 38,53 до 51,64 т/га, а в праці (Lutz N. et al. 2024) – від 5,7 до 72,4 т/га.

Розподіл за комірками растра значень AGBD, змодельованих для урбогеосистеми м. Буча за методом RF з ГР=10 м (див. табл.3), найбільш близько відповідає вельми асиметричному розподілу (Самойленко, Діброва, 2025) згідно з гістограмою рис.5. Середнє значення цього емпіричного розподілу становить 72,5 т/га. Більшість значень потрапляють у діапазон стандартного відхилення, відкладеного в обидва боки від середнього значення.

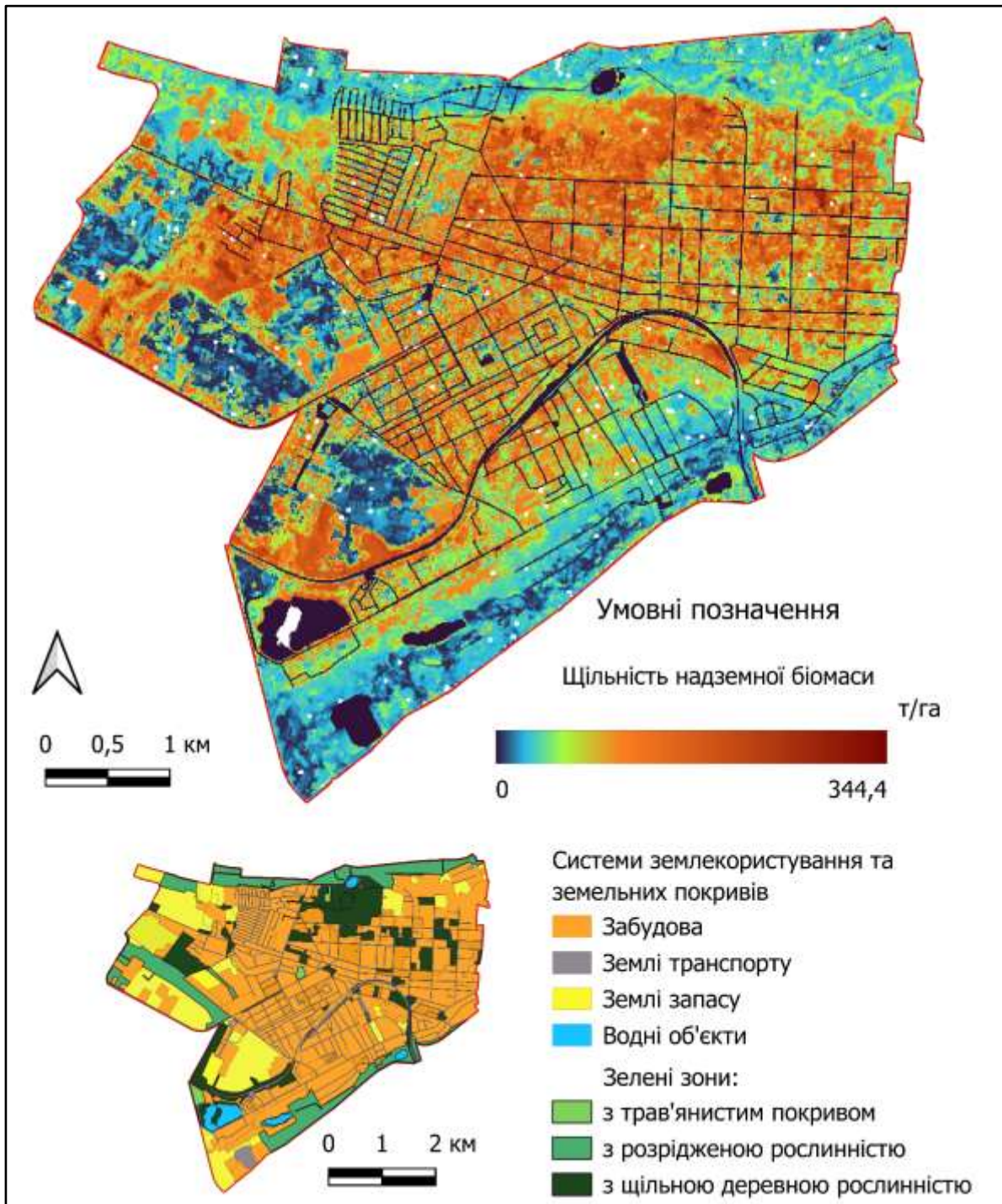


Рис. 6. Змодельований розподіл щільності надземної біомаси (AGBD) для урбогеосистеми м. Буча

Геопросторовий розподіл отриманих для досліджуваної геосистеми значень AGBD (Рис. б) дозволяє дійти наступних висновків. Максимальні обсяги надземної біомаси спостерігаються в зелених зонах з деревною рослинністю. Натомість в зелених зонах та на землях запасу із трав'янистою рослинністю, щільність надземної біомаси є дуже низькою. Тому середні значення AGBD для зелених зон у цілому не набагато перевищує середні значення у цілому для забудованих територій (82,3 т/га проти 80,5 т/га). Високі значення AGBD в межах забудованих територій свідчать про високий ступінь їхнього озеленення. Землі запасу представлено переважно покриттями з розрідженою рослинністю, тому середні значення AGBD для них є меншими (60,1 т/га).

Висновки. Проведено моделювання щільності надземної біомаси (AGBD) для малих урбогеосистем на прикладі урбогеосистеми м. Буча з використанням лідарних даних місії GEDI. Для моделювання було застосовано два методи машинного навчання (Extreme Gradient Boosting та Random Forest), а за предиктори правили цифрова модель висот (за даними SRTM та щодо нахилу поверхні та експозиції схилів), а також спектральні індекси (NDVI, EVI, NDWI, NDBI та SAVI). Була перевірена та підтверджена доцільність застосування ЦМВ як предиктора. Під час моделювання було визначено необхідність інтегрування додаткового предиктора – шару-маски, який присвоює коміркам растра, які відповідають дорогам та водним об'єктам, нульове значення AGBD, не порушуючи алгоритм моделювання для інших систем землекористування. Найкращі результати показало моделювання методом Random Forest з використанням ЦМВ, спектральних індексів та шару-маски, застосовуючи геопросторове розрізнування моделювання в 10 м.

Порівняння геопросторового розподілу AGBD з даними про системи землекористування та земельних покриттів засвідчило відносну невелику різницю в середній щільності надземної біомаси для оцінених у цілому зелених зон (82,3 т/га), житлової забудови (80,5 т/га) та земель запасу (60,1 т/га). Це зумовлене як досить високим ступенем озеленення забудови у малих урбогеосистемах, так і великою мірою усереднення змодельованих даних для укрупненої класифікації систем LULC. Перспективним напрямком розвитку зазначеного моделювання є деталізація даних про системи землекористування та земельні покриття, зокрема поділ зелених зон за типами рослинності на хвойні, мішані, широколистяні та трав'янисті, поділ забудови на житлову та нежитлову тощо, задля більш ґрунтовного дослідження закономірностей геопросторового розподілу AGBD. В подальшому також планується використати отримані модельні дані для оцінки екосистемних послуг депонування вуглецю малими урбогеосистемами.

Список використаних джерел

- Copernicus Data Space Ecosystem Browser. 2025. URL: <https://browser.dataspace.copernicus.eu>
- Copernicus Sentinel-2 (processed by ESA), 2021, MSI Level-2A BOA Reflectance Product. Collection 1. European Space Agency. https://doi.org/10.5270/S2_znk9xsi
- Dubayah, R.O., J. Armston, J.R. Kellner, L. Duncanson, S.P. Healey, P.L. Patterson, S. Hancock, H. Tang, J.M. Bruening, M.A. Hofton, J.B. Blair, & S.B. Luthcke. 2022. GEDI L4A Footprint Level Aboveground Biomass Density, Version 2.1. ORNL DAAC, Oak Ridge, Tennessee, USA. <https://doi.org/10.3334/ORNLDAAAC/2056>
- Google Colaboratory. (2025). <https://colab.research.google.com/>
- Google. (2025). Google Satellite map data.
- Gülçin, D., & van den Bosch, C.C.K. (2021). Assessment of Above-Ground Carbon Storage by Urban Trees Using LiDAR Data: The Case of a University Campus. *Forests*, 12, 62. <https://doi.org/10.3390/f12010062>
- Kandii, M., & Samoilenko, V. (2025). Remote sensing data for climate neutrality of urbogeosystems assessment using InVest model. International Conference of Young Professionals «GeoTerrace-2025», October 6–9, 2025 in the Institute of Geodesy in Lviv Polytechnic National University (Lviv, Ukraine). <https://openreviewhub.org/geoterrace/paper-2025/remote-sensing-data-climate-neutrality-urbogeosystems-assessment-using-invest>
- Lutz, N., Rodriguez-Veiga, P., & Menor, I.O. (2024). Estimating vegetation structure and aboveground carbon storage in Western Australia using GEDI LiDAR, Landsat and Sentinel data. *Environmental Research: Ecology*, 3, 045004. <https://doi.org/10.1088/2752-664X/ad7f5a>
- Liang, M., Duncanson, L., Silva, J. A., Sedano, F. (2023). Quantifying aboveground biomass dynamics from charcoal degradation in Mozambique using GEDI Lidar and Landsat. *Remote Sensing of Environment*, Vol, 284, 113367. <https://doi.org/10.1016/j.rse.2022.113367>
- Myroniuk, V., Zibtsev, S., Bogomolov, V., Goldammer, J.G., Soshenskyi, O., Levchenko, V., & Matsala, M. (2023). Combining Landsat time series and GEDI data for improved characterization of fuel types and canopy metrics in wildfire simulation, *Journal of Environmental Management*, Vol. 345, 118736. <https://doi.org/10.1016/j.jenvman.2023.118736>.
- NASA's Earth science data (2025). URL: <https://search.earthdata.nasa.gov/>
- Oak Ridge National Laboratory (ORNL) Distributed Active Archive Center (DAAC). (2025). GEDI L4A Footprint Level Aboveground Biomass Density, Version 2.1 User Guide. URL: https://daac.ornl.gov/GEDI/guides/GEDI_L4A_AGB_Density_V2.html
- OpenStreetMap contributors. (2015). Planet dump [Data file from \$date of database dump\$]
- Spracklen, B.; Spracklen, D.V. (2021). Determination of Structural Characteristics of Old-Growth Forest in Ukraine Using Spaceborne LiDAR. *Remote Sensing*, 13, 1233. <https://doi.org/10.3390/rs13071233>
- Thinley, J., & Ndehedehe, Ch. (2025). Quantifying aboveground biomass of small urban remnant forest in South-East Queensland from global ecosystem dynamic investigation laser footprints and sentinel-2 imagery. *Nature-Based Solutions*, 8, 100267. <https://doi.org/10.1016/j.nbsj.2025.100267>

- Tikuye, B.G., & Ray, R.L. (2025). Estimating aboveground biomass using environmental covariates and a machine-learning approach in the Lower Brazos River Basin, Texas, USA. *Applied Computing and Geosciences*, 27, 100289. <https://doi.org/10.1016/j.acags.2025.100289>
- United States Geological Survey. (2025). Earth Explorer. URL: <https://earthexplorer.usgs.gov/>
- Wang, C., Zhang, W., Ji, Y., Marino, A., Li, C., Wang, L., Zhao, H., & Wang, M. (2024). Estimation of Aboveground Biomass for Different Forest Types Using Data from Sentinel-1, Sentinel-2, ALOS PALSAR-2, and GEDI. *Forests*, 15, 215. <https://doi.org/10.3390/f15010215>
- Wang, J., Xiang, C., & Liang, A. (2025). Estimation of Forest Aboveground Biomass in China Based on GEDI and Sentinel-2 Data: Quantitative Analysis of Optical Remote Sensing Saturation Effect and Terrain Compensation Mechanisms. *Remote Sensing*, 17, 3437. <https://doi.org/10.3390/rs17203437>
- Wang, Y., Wang, H., Wang, C., Zhang, S., Wang, R., Wang, S., & Duan, J. (2024). Co-Kriging-Guided Interpolation for Mapping Forest Aboveground Biomass by Integrating Global Ecosystem Dynamics Investigation and Sentinel-2 Data. *Remote Sensing*, 16, 2913. <https://doi.org/10.3390/rs16162913>
- Zhao, X., Hu, W., Han, J., Wei, W., & Xu, J. (2024). Urban Above-Ground Biomass Estimation Using GEDI Laser Data and Optical Remote Sensing Images. *Remote Sensing*, 16, 1229. <https://doi.org/10.3390/rs16071229>
- Shpylevskiy, I.I., Tokar V.O., Hubenko T.V., Mukha V.H., Pustotin O.S., Buchok V.S. (2015). Heneralnyi plan m. Bucha. Plan zonuvannya (zoninh) m. Bucha Kyivskoi oblasti. [General plan of Bucha. Zoning plan of Bucha, Kyiv Region] DP Ukrainyskyi derzhavnyi naukovo-doslidnyi instytut proektuvannya mist "Dipromisto" imeni Yu.M.Bilokonnia. (In Ukrainian) URL: <https://bucha-rada.gov.ua/content/viddil-mistobuduvannya-ta-arhitektury> (Date: 30.09.2025 p.)
- Samoilenko, V.M., Dibrova, I.O. (2025). Imovirnisni metody v heoekolohii ta analizi servisiv dovkillia : pidruchnyk-monohrafiia (elektronna versia). [Probabilistic methods in geoecology and analysis of environmental services] Kyiv: Nika-Tsentr. 372 p. (In Ukrainian) https://www.researchgate.net/publication/391300672_Samoilenko_VM_Dibrova_IO_Imovirnisni_metodi_v_geoekologii_ta_analizi_servisiv_dovkilla_Pidruchnik-monografia_Kiiv_Nika-Centr_2025

Отримано редакцією журналу / Received: 15.01.26
Прорецензовано / Revised: 20.02.26
Схвалено до друку / Accepted: 27.02.26