SCIENTIFIC AND TECHNICAL JOURNAL MACHINE BUILDING

Artificial intelligence and geospatial technologies for biodiversity hotspot research: a Scopus-based bibliometric review and implications for Uzbekistan

Berdiyeva Zarina Farkhod qizi
Department of Industrial Ecology,
Bukhara State Technical University, Bukhara, Uzbekistan
e-mail: zarinaberdiyeva74@gmail.com

ABSTRACT

The convergence of artificial intelligence (AI), machine learning and geospatial technologies has transformed how ecologists detect, map and monitor biodiversity in threatened landscapes. This review synthesizes 82 peer-reviewed articles retrieved from Scopus (2013-2025) applying AI, GIS and remote-sensing-based methods to biodiversity hotspot research. A PRISMA-adapted procedure reduced an initial 800 records to 82 relevant documents, 82 of which were fully analyzed. Annual output expanded sharply after 2018, with most documents published in 2023-2025. Random Forest was the dominant algorithm (34 of 82 documents), followed by species-distribution-modelling frameworks (20), Sentinel-1/2 imagery (18) and deep learning (15). The literature clusters around machine-learning classification, deep-learning survey interpretation, UAV/hyperspectral/LiDAR platforms, satellite monitoring, eDNA and citizen-science integration, and invasive-species detection. The United States, India and Brazil were the most cited producing countries; Remote Sensing, Ecological Informatics and Ecological Indicators were the leading outlets. No study addressed Uzbekistan or Central Asia, revealing a clear geographical gap, although transferable evidence from arid, steppe and grassland ecosystems elsewhere is identified. AI-assisted geospatial monitoring is methodologically mature for operational use, but its transfer to data-scarce Central Asian contexts requires dedicated empirical validation, capacity building and open-data infrastructure.

Keywords: artificial intelligence, machine learning, Random Forest, remote sensing, geographic information systems, biodiversity hotspot, species distribution modelling, Uzbekistan

Искусственный интеллект и геопространственные технологии в исследованиях горячих точек биоразнообразия: библиометрический обзор на основе Scopus и рекомендации для Узбекистана

Сочетание искусственного интеллекта (ИИ), машинного обучения и геопространственных технологий коренным образом изменило подходы экологов к выявлению, картированию и мониторингу биоразнообразия в находящихся под угрозой

Ключевые слова: искусственный интеллект, машинное

исчезновения ландшафтах. В данном обзоре обобщены результаты 82 рецензируемых научных статей, отобранных из базы данных Scopus за 2013–2025 гг., в которых применялись методы на основе искусственного интеллекта, ГИС и дистанционного зондирования для исследования «горячих точек» биоразнообразия. Адаптированная процедура PRISMA позволила сократить первоначальную выборку из 800 публикаций до 82 релевантных документов, все из которых были под-вергнуты полному анализу. После 2018 года наблюдался резкий рост ежегодного количе-ства публикаций, причём большинство документов было опубликовано в 2023–2025 гг. Наиболее распространённым алгоритмом оказался Random Forest, использованный в 34 из 82 документов, за ним следуют методы моделирования распространения видов (20), данные изображений Sentinel-1/2 (18) и глубокое обучение (15). Научная литература в основном сосредоточена на классификации с использованием машинного обучения, интерпретации данных наблюдений посредством глубокого обучения, применении UAV/гиперспектральных/LiDAR-платформ, спутниковом мониторинге, интеграции eDNA и гражданской науки, а также выявлении инвазивных видов. США, Индия и Бразилия явля-лись наиболее цитируемыми странами по количеству опубликованных исследований; ведущими научными изданиями выступили Remote Sensing, Ecological Informatics и Ecological Indicators. Ни одно из рассмотренных исследований не было посвящено Узбекистану или Центральной Азии, что свидетельствует о наличии очевидного географического пробела. Вместе с тем были выявлены результаты исследований засушливых, степных и луговых экосистем других регионов, которые могут быть адаптированы к условиям Центральной Азии. Геопространственный мониторинг с использованием искусственного интеллекта достиг методологической зрелости, необходимой для практического применения, однако его перенос в условия Центральной Азии, характеризующиеся дефицитом данных, требует проведения специализированной эмпирической валидации, развития кадрового потенциала и создания инфраструктуры открытых данных.

обучение, Random Forest, дистанционное зондирование, географические информационные системы, «горячие точки» биоразнообразия, моделирование распространения видов, Узбекистан.

Biologik xilma-xillik markazlarini tadqiq etishda sun’iy intellekt va geomakon texnologiyalari: Scopus ma’lumotlar bazasiga asoslangan bibliometrik tahlil va O‘zbekiston uchun tavsiyalar

ANNOTATSIYA

Sun’iy intellekt (SI), mashinali o‘qitish va geofazoviy texnologiyalarning uyg‘unlashuvi ekologlarning xavf ostidagi landshaftlarda biologik xilma-xillikni aniqlash, xaritalash va monitoring qilish usullarini tubdan o‘zgartirdi. Mazkur sharh Scopus ma’lumotlar bazasidan (2013–2025-yillar) olingan, biologik xilma-xillikning «qaynoq nuqtalari»ni tadqiq etishda sun’iy intellekt, GIS va masofadan zondlashga asoslangan usullarni qo‘llagan 82 ta taqrizdan o‘tgan ilmiy maqolani umumlashtiradi. PRISMA

Kalit so‘zlar: sun’iy intellekt, mashinali o‘qitish, Random Forest, masofadan zondlash, geografik axborot tizimlari, biologik xilma-xillikning

metodologiyasiga moslashtirilgan tartib asosida dastlabki 800 ta manbaadan 82 tasi tegishli hujjat sifatida saralandi va ularning barchasi to‘liq tahlil qilindi. Yillik ilmiy nashrlar soni 2018-yildan keyin keskin oshgan bo‘lib, hujjatlarning asosiy qismi 2023–2025-yillarda chop etilgan. Random Forest algoritmi eng ko‘p qo‘llanilgan usul bo‘lib, 82 ta hujjatning 34 tasida ishlatilgan. Undan keyingi o‘rinlarda turlarning tarqalishini modellashtirish (20 ta), Sentinel-1/2 tasvirlari (18 ta) va chuqur o‘qitish (15 ta) turadi. Ilmiy adabiyot-larda asosiy yo‘nalishlar mashinali o‘qitish asosidagi klassifikatsiya, chuqur o‘qitish yordamida kuzatuv ma’lumotlarini talqin qilish, UAV/giperspektral/LiDAR platformalari, sun’iy yo‘ldosh monitoringi, eDNA va fuqarolik fanini integratsiyalash hamda invaziv turlarni aniqlashdan iborat. Eng ko‘p iqtibos keltirilgan tadqiqotlarni nashr etgan mamlakatlar AQSh, Hindiston va Braziliya bo‘lgan; Remote Sensing, Ecological Informatics va Ecological Indicators jurnallari yetakchi nashrlar sifatida qayd etilgan. Hech bir tadqiqot O‘zbekiston yoki Markaziy Osiyoga bag‘ishlanmagan bo‘lib, bu aniq geografik bo‘shliq mavjudligini ko‘rsatadi. Shu bilan birga, boshqa hududlardagi qurg‘oqchil, dasht va o‘tloq ekotizimlarida olingan hamda Markaziy Osiyo sharoitiga moslashtirilishi mumkin bo‘lgan ilmiy dalillar aniqlandi. Sun’iy intellekt yordamidagi geofazoviy monitoring amaliy qo‘llash uchun metodologik jihatdan yetarlicha rivojlangan, biroq uni ma’lumotlar tanqis bo‘lgan Markaziy Osiyo sharoitiga joriy etish maxsus empirik validatsiya, mutaxassislar salohiyatini oshirish va ochiq ma’lumotlar infratuzilmasini rivojlantirishni talab etadi.

«qaynoq nuqtalari», turlarning tarqalishini modellashtirish, O‘zbekiston.

1. Introduction

Global biodiversity loss is spatially uneven, concentrated in hotspots under intense anthropogenic pressure, where traditional field surveys are too labour-intensive to match the resolution conservation planning requires. This gap has been addressed by the convergence of GIS, remote sensing (Sentinel, Landsat, UAV, hyperspectral and LiDAR platforms) and artificial intelligence-especially Random Forest-which together convert spatial layers into maps of species occurrence and habitat suitability [1,14], and increasingly incorporate eDNA metabarcoding [4,5,9,20], citizen-science records [2,7,8] and deep learning for image-based detection [12].

Despite this maturation, AI-assisted biodiversity research remains geographically uneven: citation impact concentrates in the United States, India, Brazil, the United Kingdom and Australia, while Central Asia’s arid, steppe and mountain ecosystems are essentially absent from the indexed literature. Uzbekistan spans desert, steppe, riparian (tugai) and mountain ecosystems of regional significance, yet no AI- or GIS-based biodiversity assessment of the country was found in the corpus reviewed here, even though toolkits demonstrated in analogous systems-the desert steppe of China [15] and the semi-arid rangelands of Kenya [11]-are, in principle, transferable. This review therefore (i) characterises the growth, geographic distribution and thematic composition of the 2013-2025 AI/GIS biodiversity-hotspot literature; (ii) synthesises the evidence into thematic clusters; and (iii) evaluates its relevance to Uzbekistan.

2. Materials and Methods

This review is based on a bibliographic corpus retrieved from Scopus, restricted to 2010-2025 and English-language documents, combining biodiversity-, AI- and geospatial-technology-related search terms applied to title, abstract and keyword fields (800 initial records).

Document selection followed a PRISMA-adapted workflow (Figure 1): 240 duplicates removed (560 remaining); 160 non-Article/Review records excluded (400 remaining); restriction to four relevant Scopus subject areas-Environmental Science, Earth and Planetary Sciences, Computer Science, Agricultural and Biological Sciences (352 remaining); and an 80% relevance screening of abstracts, excluding 270 further. Eighty-two records satisfied the final criteria, 82 of which had complete metadata and form the corpus analysed here.

The 82 documents were processed with Bibliometrix/Biblioshiny and VOSviewer to characterise annual production, leading sources/countries, and keyword co-occurrence, and were thematically classified by algorithm, sensor platform and ecosystem type using the original authors' own terminology. No new experimental analysis was generated; all figures and statistics below derive from the bibliometric analysis of the corpus or the reported content of the reviewed documents.

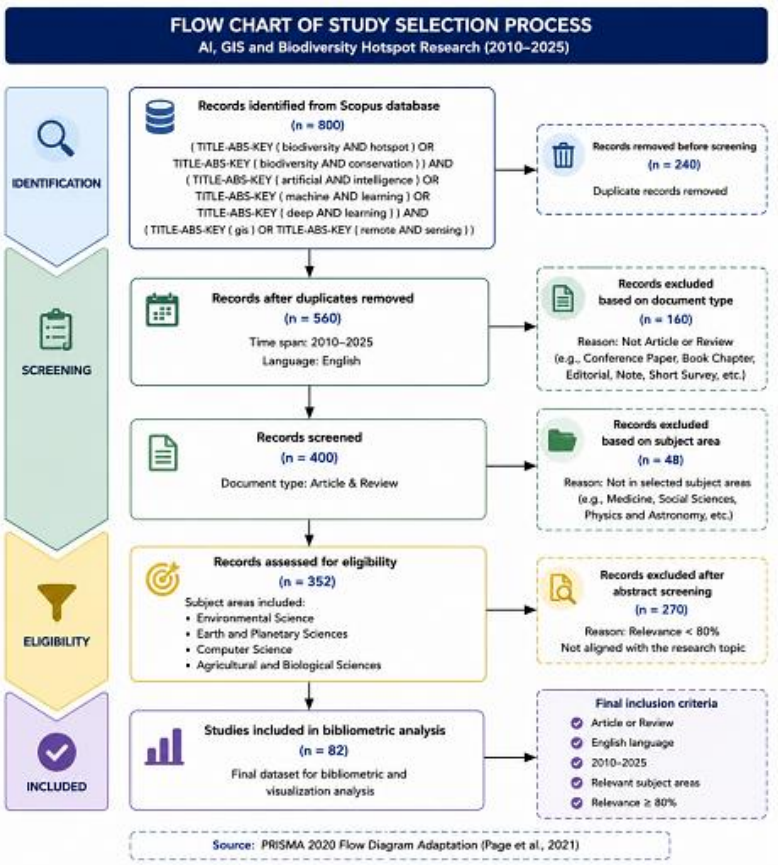

Figure 1. PRISMA-adapted flow diagram of the study selection process for the AI/GIS/biodiversity-hotspot literature corpus (2010-2025), from initial identification (n = 800) through screening, eligibility and final inclusion (n = 82).

3. Results

3.1 Bibliometric overview

The 82 documents were published between 2013 and 2025, with output rising sharply after 2018-eleven documents in 2023, fifteen in 2024 and twenty-nine in 2025, more than half the corpus in the final two years. Of the 82 records, 73 were research articles and seven reviews. Leading outlets were Remote Sensing (12), Ecological Informatics (6) and Ecological Indicators (5). The United States (134 citations), India (115) and Brazil (114) were the most cited producing countries, followed by the United Kingdom, Australia, Germany, Israel, Italy, Switzerland and South Africa (Figure 2). No document originated from, or was co-authored by an institution in, Uzbekistan or another Central Asian country.

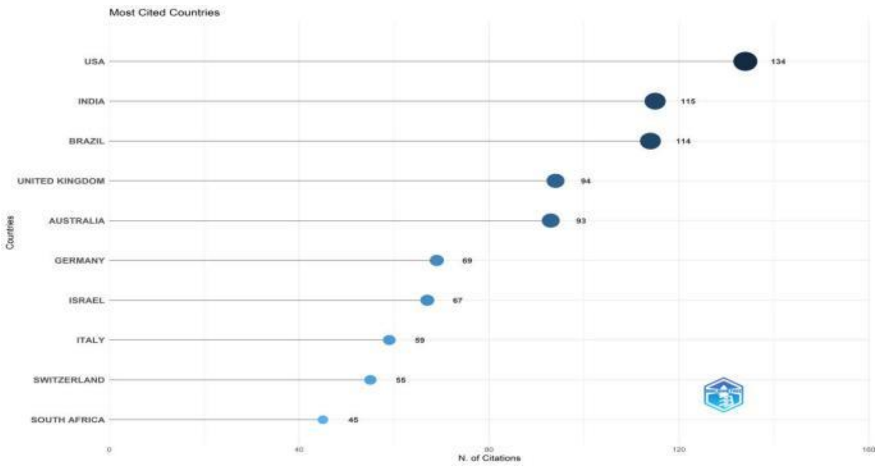

Figure 2. Most cited countries in the reviewed corpus, by total citations received (Bibliometrix/Biblioshiny analysis).

3.2 Algorithms, sensor platforms and thematic clusters

Random Forest was the dominant algorithm (34 of 82 documents), applied from prediction of tree-species distribution in data-poor forests [1] to diversity-index prediction across China's desert steppe [15], and consistently favoured for its robustness to noisy, high-dimensional predictor sets [1,14,18]. Deep learning (15 documents) offered advantages mainly with large, spectrally complex datasets - UAV hyperspectral imagery or LiDAR [12] - though its computational demands limit use in data-scarce settings. Satellite remote sensing, led by Sentinel-1/2 (18 documents), and UAV/hyperspectral/LiDAR platforms formed the observational backbone, complemented by cloud-based platforms such as Google Earth Engine [19].

Species-distribution modelling recurred across 20 documents, spanning taxa from red panda and black rhino to sub-alpine grassland species [10]. eDNA metabarcoding was integrated with remote-sensing predictors for fish and arthropod biodiversity [3,4,5,9,20], and citizen-science data were combined with machine-learning predictors to monitor waterbird habitats, range-limit species and protected areas [2,7,8]. Invasive-species detection-Sentinel-2 classification of shrub encroachment in Kenya [11] and hyperspectral discrimination in the Western Ghats [3] - and grassland/mountain biodiversity under environmental change, including the ORCHAMP network [16], were further recurring clusters directly relevant to Uzbekistan's steppe and mountain biomes.

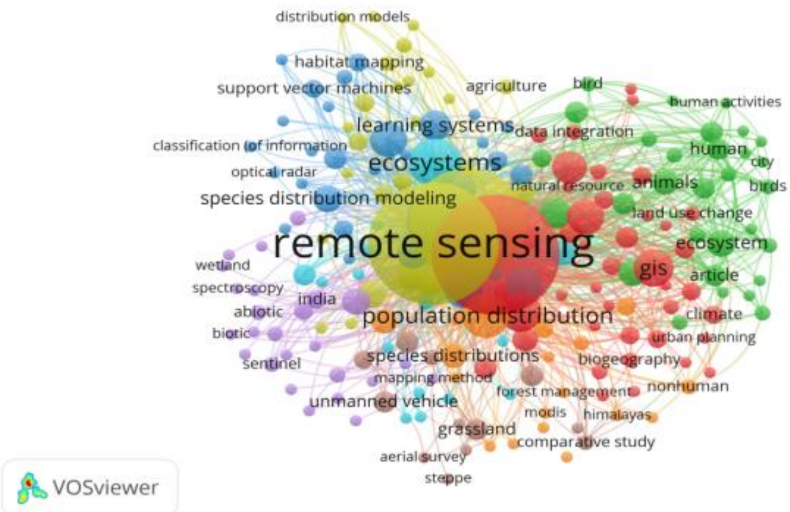

Figure 3. Keyword co-occurrence network (VOSviewer) of author keywords across the 82-document corpus.

4. Discussion

The reviewed evidence shows strong convergence around Random Forest, reflecting its robustness to noisy, high-dimensional predictor sets [1,14,18], complemented by deep learning where large, high-dimensional imagery is available [12]. Sensor platforms present a resolution–extent trade-off-satellite time series (Sentinel-1/2) for landscape-scale coverage versus UAV/hyperspectral/LiDAR for fine-scale structure-while eDNA and citizen-science data are increasingly integrated as complementary streams [2,4,5,7,8,9,20]. Several limitations, reported by the reviewed authors themselves, deserve note: model transferability across regions is not guaranteed without local recalibration [1,18]; ground-truth data scarcity, rather than algorithm choice, is the primary bottleneck limiting accuracy [17]; and gaps in standardisation of biodiversity variables complicate cross-study comparison [13,17]. These point toward the investments-local training data, regional calibration, standardised protocols-needed to extend this evidence base to regions such as Central Asia.

5. Recommendations for Uzbekistan

The absence of any AI-, GIS- or remote-sensing-based biodiversity hotspot study addressing Uzbekistan or Central Asia in the reviewed corpus is itself the review's most direct finding for the country; any recommendation must therefore be framed as an inference of transferability from analogous systems studied elsewhere, not as a synthesis of Uzbekistan-specific findings.

Uzbekistan's desert and steppe ecosystems are comparable to the desert-step-pe system validated in China [15] and the semi-arid rangelands mapped with Sentinel-2 ensemble classifiers in Kenya [11]; combining Sentinel-1/2 time series with Random Forest is, in principle, transferable to these zones, subject to local ground-truth data [1,14]. Uzbekistan's mountain systems (western Tian Shan, Pamir-Alai) are similarly comparable to sub-alpine systems studied via Random Forest-based richness prediction [10] and to networks such as ORCHAMP [16]. Two infrastructural enablers could lower the barrier to adoption: cloud-based platforms such as Google Earth Engine [19], and citizen-science/eDNA integration [2,4,5,7,8,9,20], both of which reduce dependence on local high-performance computing and systematic field survey alone.

On this basis: methodological transfer of Sentinel-1/2 and Random Forest-based protocols, validated in analogous desert-steppe and rangeland systems [11,15], to Uzbekistan's steppe and desert biomes; adoption of cloud-based processing platforms [19]; and recognition that the absence of any reviewed Central Asian case study is a well-defined research gap whose first empirical closure would be of national and international relevance.

6. Conclusions

This review synthesised 82 peer-reviewed documents retrieved through a PRISMA-adapted Scopus search to characterise AI and geospatial-technology-based biodiversity-hotspot research (2013-2025). The literature shows accelerating growth and strong convergence around Random Forest, complemented by deep learning for high-dimensional imagery, with Sentinel-1/2 as the backbone of landscape-scale monitoring and eDNA/citizen-science data increasingly integrated as complementary streams. At the same time, the review documents a clear geographic gap: no study addresses Uzbekistan or Central Asia. Given the maturity of Random Forest- and Sentinel-based approaches in analogous desert, steppe and mountain systems elsewhere, this gap reflects an absence of documented application rather than a lack of methodological readiness, and closing it will require local ground-truth data collection, regional model calibration, and cloud-based processing infrastructure.

References

1. Ahmadi K.; Mahmoodi S.; Pal S.C.; et al. (2023). Improving species distribution models for dominant trees in climate data-poor forests using high-resolution remote sensing. Ecological Modelling, 475.
2. Antonelli A.; Dhanjal-Adams K.L.; Silvestro D. (2023). Integrating machine learning, remote sensing and citizen science to create an early warning system for biodiversity. Plants People Planet, 5(3), 307-316.
3. Arasumani M.; Singh A.; Bunyan M.; Robin V.V. (2021). Testing the efficacy of hyperspectral (AVIRIS-NG), multispectral (Sentinel-2) and radar (Sentinel-1) remote sensing images to detect native and invasive non-native trees. Biological Invasions, 23(9), 2863-2879.

4. Bauknecht R.; Pellissier L.; Brosse S.; et al. (2025). Combining environmental DNA and remote sensing variables to model fish biodiversity in tropical river ecosystems. Ecological Informatics, 90.
5. Bizzozzero M.R.; Marfurt S.M.; Altermatt F.; et al. (2025). Integrating Environmental DNA Metabarcoding and Remote Sensing Reveals Known and Novel Fish Diversity Hotspots in a World Heritage Area. Diversity and Distributions, 31(11).
6. Cheng K.; Zhang C.; Fan Y.; Diao H.; Xia S. (2025). AI-based approaches for wetland mapping and classification: A review of current practices and future perspectives. Artificial Intelligence in Geosciences, 6(2).
7. Jiang X.; Zhang L.; Feng K.; et al. (2025). Mapping urban waterbird habitats using machine learning and citizen science: A multi-scale analysis in Brussels, Belgium. Journal of Environmental Management, 395.
8. Krämer A.; Meyer H.; Buchholz S. (2025). Habitat Suitability of the Sand Lizard (Lacerta agilis) at Its Distribution Limit — An Analysis Based on Citizen Science Data and Machine Learning. Journal of Biogeography, 52(5).
9. Li Y.; Devenish C.; Tosa M.I.; et al. (2024). Combining environmental DNA and remote sensing for efficient, fine-scale mapping of arthropod biodiversity. Philosophical Transactions of the Royal Society B: Biological Sciences, 379(1904).
10. Mashiane K.; Ramoelo A.; Adelabu S. (2024). Prediction of species richness and diversity in sub-alpine grasslands using satellite remote sensing and random forest machine-learning algorithm. Applied Vegetation Science, 27(2).
11. Muthoka J.M.; Salakpi E.E.; Ouko E.; et al. (2021). Mapping Opuntia stricta in the arid and semi-arid environment of Kenya using Sentinel-2 imagery and ensemble machine learning classifiers. Remote Sensing, 13(8).
12. Pettorelli N.; Williams J.; Schulte to Bühne H.; Crowson M. (2025). Deep learning and satellite remote sensing for biodiversity monitoring and conservation. Remote Sensing in Ecology and Conservation, 11(2), 123-132.
13. Reddy C.S.; Athira K.; Mudalkar S. (2025). Evolution of Global Biodiversity Monitoring: Innovations, Modern Approaches and Integration with Essential Variables. Journal of the Indian Society of Remote Sensing, 53(7), 2017-2030.
14. Satish K.V.; Dugesar V.; Pandey M.K.; et al. (2023). Seeing from space makes sense: Novel earth observation variables accurately map species distributions over Himalaya. Journal of Environmental Management, 325.
15. Tang Z.; Xuan C.; Zhang T.; et al. (2025). A Novel Remote Sensing Framework Integrating Geostatistical Methods and Machine Learning for Spatial Prediction of Diversity Indices in the Desert Steppe. Agriculture (Switzerland), 15(18).
16. Thuiller W.; Saillard A.; Abdulhak S.; et al. (2024). ORCHAMP: an observation network for monitoring biodiversity and ecosystem functioning across space and time in mountainous regions. Comptes Rendus - Biologies, 347, 223-247.
17. Timmermans J.; Daniel Kissling W. (2023). Advancing terrestrial biodiversity monitoring with satellite remote sensing in the context of the Kunming-Montreal global biodiversity framework. Ecological Indicators, 154.
18. Venter Z.S.; Roos R.E.; Nowell M.S.; et al. (2023). Comparing Global Sentinel-2 Land Cover Maps for Regional Species Distribution Modeling. Remote Sensing, 15(7).
19. Zhidebayeva A.; Syrlybekkyzy S.; Taizhanova L.; et al. (2025). Emerging frontiers in ecosystem and biodiversity monitoring using remote sensing and geographic information systems. Global Journal of Environmental Science and Management, 11(4), 1791-1818.
20. Zong S.; Brantschen J.; Zhang X.; et al. (2024). Combining environmental DNA with remote sensing variables to map fish species distributions along a large river. Remote Sensing in Ecology and Conservation, 10(2), 220-235.