Overview of GeoLifeCLEF 2025: Plant Species Presence Prediction with Environmental and High-resolution Remote Sensing Data

Lukas Picek[1,2], César Leblanc[1,3], Théo Larcher[1], Maximilien Servajean[4], Pierre Bonnet[3] and Alexis Joly[1]

[1]INRIA, LIRMM, Univ Montpellier, CNRS, Montpellier, France
[2]Department of Cybernetics, FAV, University of West Bohemia in Pilsen, Czechia
[3]AMAP, Univ Montpellier, CIRAD, CNRS, INRAE, IRD, Montpellier, France
[4]LIRMM, AMIS, Univ Paul Valéry Montpellier, Univ Montpellier, CNRS, France

Abstract

GeoLifeCLEF 2025 competition, organized as part of the LifeCLEF and FGVC workshops, challenges participants to predict plant species composition at high spatial resolution across Europe using multimodal environmental data. The task builds on a large-scale dataset that combines 5 million Presence-Only (PO) observations and approximately 100,000 standardized Presence–Absence (PA) surveys, paired with Sentinel-2 imagery, Landsat time series, climate rasters, and soil descriptors. This year's edition introduced two major challenges: a geographically shifted test set with plots from previously unseen regions with different species distribution, thereby including many rare species that are under-reported by citizen scientists. These changes increased the modeling difficulty and emphasized the need for generalization under spatial shift and class imbalance. In this paper, we summarize the task design, dataset characteristics, evaluation protocol, participant approaches, and competition results, and discuss implications for scalable species distribution modeling and biodiversity monitoring.

Keywords

LifeCLEF, biodiversity, environmental data, species distribution, prediction, evaluation, benchmark, methods comparison, presence-only data, presence-absence, model performance, remote sensing

1. Introduction

Monitoring plant species distributions at high spatial resolution is essential for understanding ecosystem dynamics and informing conservation efforts. However, collecting standardized species observations over large areas remains resource-intensive and geographically constrained. Species distribution models (SDMs) offer a scalable solution by learning to predict species presence from a combination of species occurrence data and environmental predictors. These occurrence data include Presence-Absence (PA) records, which systematically document whether a species is detected or not at surveyed locations, and Presence-Only (PO) records, which opportunistically record only where a species has been observed, without information on absences.

In recent years, deep learning–based SDMs (deep-SDMs) have demonstrated improved accuracy by leveraging heterogeneous environmental data sources, including multi-spectral satellite imagery, climatic time series, and edaphic (soil) variables [1, 2, 3]. Despite these advances, several challenges remain. PO data (available at large scale from platforms such as Pl@ntNet and iNaturalist) are relatively sparse, spatially extensive, and subject to sampling bias and annotation noise [4, 5, 6].

In contrast, standardized PA data are less affected by labeling noise but are geographically concentrated in a few well-sampled regions. Additionally, plant species distributions are highly imbalanced, with most taxa being rare, making model training under limited supervision difficult [7]. Finally, environmental inputs vary widely in terms of resolution, format, and temporal depth, requiring models that can integrate multi-source, multi-scale data.

CLEF 2025 Working Notes, 9 – 12 September 2025, Madrid, Spain
lukas.picek@inria.fr (L. Picek); cesar.leblanc@inria.fr (C. Leblanc); alexis.joly@inria.fr (A. Joly)
0000-0002-6041-9722 (L. Picek); 0000-0002-5682-8179 (C. Leblanc); 0000-0002-2161-9940 (A. Joly)

Despite these limitations, the increasing availability of multimodal environmental data and large-scale biodiversity observations opens opportunities to evaluate and improve SDMs in realistic settings. To support this, the GeoLifeCLEF challenge [8, 9, 10, 11, 12, 13] was created as part of the LifeCLEF [14, 15, 16, 17] and FGVC workshop series. Its objective is to benchmark SDMs under operational constraints, such as label imbalance, biased sampling, and spatial generalization, while promoting reproducible, scalable modeling approaches grounded in real ecological data.

The 2025 edition continues this effort by focusing on multi-species prediction across geolocated vegetation plots in Europe. Participants were tasked with predicting plant species composition using high-resolution Sentinel-2 imagery, Landsat time series, climate variables, and edaphic predictors. The training set combines approximately 90,000 PA surveys from the European Vegetation Archive (EVA) [18] and over 5 million PO observations from GBIF. Each plot is represented by multimodal environmental descriptors with variable spatial resolution, ranging from 10 m to 1 km.

This year's edition introduces two key challenges. First, the test set includes more than 14,000 test vegetation plots primarily sampled from regions not represented in the PA training data, resulting in a strong spatial distribution shift. Second, the label space includes a larger proportion of rare species, increasing the difficulty of generalization under limited supervision. Together, these conditions reflect real-world limitations of field survey coverage and taxonomic imbalance, making the task a more realistic benchmark compared to prior SDM benchmarks that often rely on data with more uniform geographic sampling and less representation of rare species.

2. Dataset and Evaluation Protocol

The dataset for GeoLifeCLEF 2025 is built directly upon the GeoPlant dataset [19], which was used in the 2024 edition [20]. The training occurrence data remains the same and includes ∼5M PO observations from GBIF and related repositories and ∼100K standardized PA surveys from EVA, covering roughly 10K species. The dataset continues to provide multimodal inputs, including: (i) Sentinel-2 image patches (RGB+NIR, 64×64, 10 meter resolution), (ii) Landsat-based satellite time series (6 spectral bands, spanning 84 seasons from 2000–2020), (iii) Monthly climatic time series (CHELSA, 2000–2019), and (iv) Raster-derived scalar and spatial predictors (e.g., elevation, land cover, soil, bioclimatic variables, human footprint). However, several notable updates and improvements have been introduced:

1. A new set of significantly **more detailed human footprint rasters** was added, now at a 30-meter resolution (compared to 1 km in previous editions). Derived from OpenStreetMap data (2021), these updated rasters capture fine-grained features such as roads, railways, and built environments with greater spatial and temporal precision (see Figure 1) than the previously used global rasters (e.g., Venter et al. 2016 [21]).

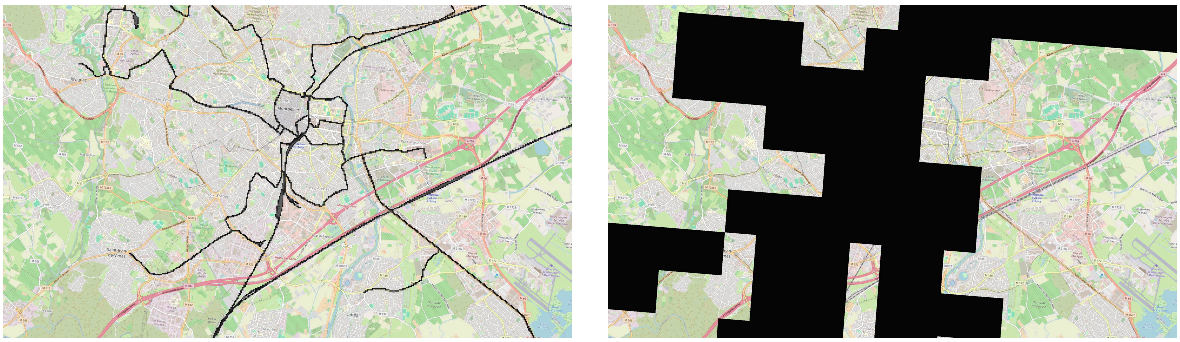

Figure 1: Upgrade in spatial detail of railway density. *Right*: The 2024 edition, with 1 km resolution. *Left*: new rasters with 30 m resolution based on 2021 OpenStreetMap data. The new rasters capture anthropogenic features with higher resolution, supporting more accurate modeling of human influence on species distributions.

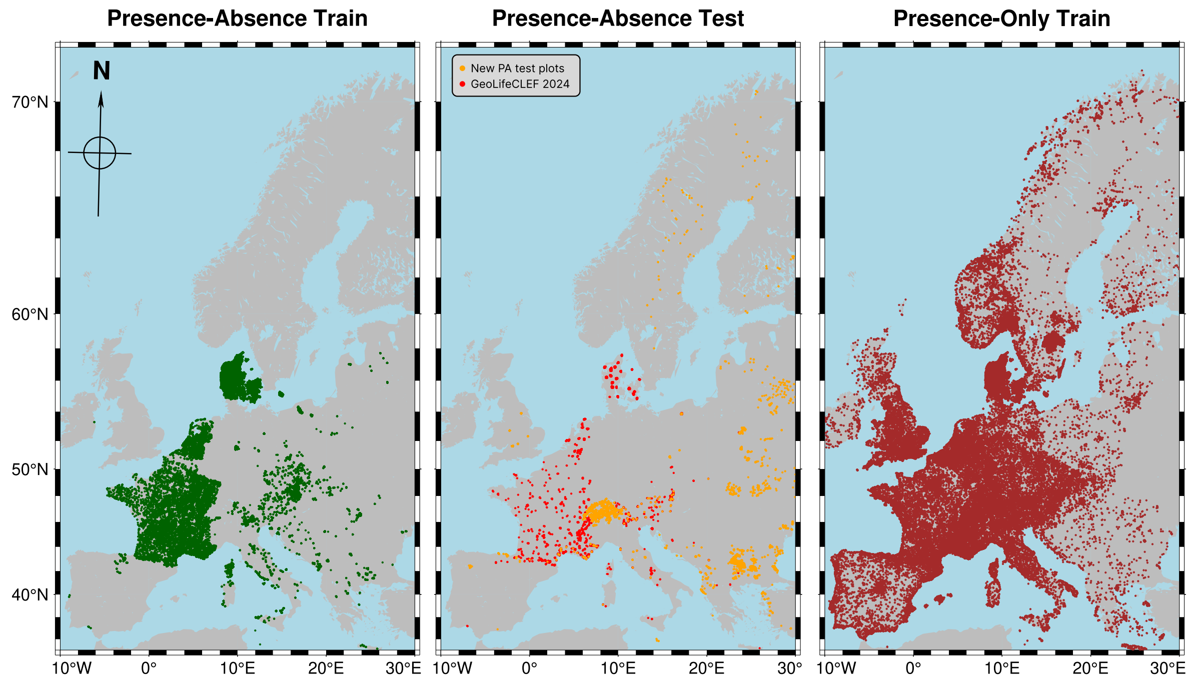

Figure 2: Geospatial coverage of the GeoLifeCLEF 2025 dataset. The Presence-Absence (PA) training data and test sites from GeoLifeCLEF 2024 are primarily concentrated in Western and Central Europe, including France, Denmark, Switzerland, Czechia, and Italy. In contrast, the new test sites in 2025 extend into previously unseen regions, particularly in Eastern and Southeastern Europe, thereby introducing a significant spatial distribution shift relative to the training data. Presence-Only (PO) training data spans the majority of habitable Europe, providing broad spatial context.

2. The last year **test set was enriched with more than 9,000 surveys** from new geographical origins (i.e., eastern and northern Europe), allowing to test geospatial generalization (see Figure 2).
3. The **SoilGrids data**, which was incorrectly exported in the 2024 dataset, **was corrected** and re-extracted. In the previous version. The land cover and soil features contained identical, non-informative values due to a processing error. This significantly reduced their utility for species prediction.
4. The **Sentinel-2 satellite data underwent a major upgrade** in format and processing. Instead of the previously used compressed JPEG images, this year's edition provides raw multi-band TIFF files, significantly improving radiometric fidelity and spatial integrity for geospatial modeling. These TIFFs include all four bands (RGB + NIR) at 10-meter resolution. Updated preprocessing and normalization techniques were provided in official tutorial notebooks, enabling more accurate and flexible use of the remote sensing inputs.

2.1. Evaluation Metric

As in the previous editions [13, 20], we use the sample-averaged F_1-score (F_1^s) as the main evaluation metric. The F_1^s measures the degree of agreement between the predicted and actual species composition observed within a specific geographical area and timeframe. In the context of ecological surveys, such as those conducted in protected areas, each survey instance i is associated with a ground-truth set of labels Y_i, representing the plant species found by experts within a defined grid. Given this setup, and a list of predicted labels $\widehat{Y}_{i,1}, \widehat{Y}_{i,2}, \ldots, \widehat{Y}_{i,R_i}$, the F_1^s can be computed by averaging the per-instance F_1 scores over all samples. Let N denote the total number of evaluation samples, then the F_1^s is computed as follows:

$$
\mathrm{F}_1^s = \frac{1}{N}\sum_{i=1}^{N}\frac{2\cdot \mathrm{TP}_i}{2\cdot \mathrm{TP}_i+\mathrm{FP}_i+\mathrm{FN}_i}, \quad \text{where} \begin{cases} \mathrm{TP}_i - \text{correctly predicted, i.e., } |\widehat{Y}_i \cap Y_i|. \\ \mathrm{FP}_i - \text{predicted but not observed, i.e., } |\widehat{Y}_i \setminus Y_i|. \\ \mathrm{FN}_i - \text{not predicted but present, i.e., } |Y_i \setminus \widehat{Y}_i|. \end{cases} \tag{1}
$$

This formulation encapsulates the precision and recall elements crucial for assessing the accuracy of predictive models in ecological studies.

2.2. Baselines

This year, we provided the same set of baselines as in the 2024 edition, covering multiple modalities. All baselines were trained exclusively on Presence–Absence (PA) data and released as executable Kaggle notebooks, complete with training and inference code. The baselines include:

1. **Naive frequency-based baselines.** This model ranked species by their frequency in the PA training data, either globally or within administrative or biogeographic regions, and it served as a simple lower bound. While this approach achieved a sample-averaged F_1^s of 0.20 in 2024, it performed poorly in 2025 (0.08), reflecting the impact of a shift in spatial distribution.
2. **CNN for bioclimatic and Landsat time series.** These baselines use 3D convolutional networks derived from ResNet-18 [22] to process time series cubes: 19×12×4 for bioclimatic data and 21×4×6 for Landsat. They provide efficient, modality-specific baselines, with sample-averaged F_1^s scores of 0.12779 (Bioclim) and 0.14415 (Landsat).
3. **CNN for Sentinel-2 imagery.** Unlike last year's Swin-v2-t baseline, the 2025 edition uses a lightweight ResNet-18 backbone to process Sentinel-2 patches (32×32, 4 channels: R,G,B + NIR). This change simplifies the model and aligns it with the other single-modality baselines. The sample-averaged F_1^s score for this baseline was 0.12213.
4. **Multimodal fusion model.** A simple MLP-based model that combines the outputs of the three ResNet-18 backbones (bioclimatic, Landsat, and Sentinel-2), illustrating the performance gains from integrating multiple environmental data sources.

The only baseline modification this year is related to Sentinel-2 preprocessing. A new notebook was released for exploratory analysis and standardized normalization. It includes per-band statistics, handling of missing pixels, and min–max scaling across all four channels. This preprocessing was also provided in the updated Sentinel-2 baseline model and separated data processing notebook. For full model architecture and training configurations, we refer the reader to the GeoLifeCLEF 2024 overview paper [20].

3. Competition Results

GeoLifeCLEF 2025 drew 41 participating teams, submitting a total of 750 entries. The final leaderboard, computed on approximately 77% of the test set, revealed a substantial drop in absolute performance compared to the last year. Of all the teams, *just* 17 teams outperformed the best provided baselines on the private leaderboard. The top-performing team, *webmaking* [23], achieved an F_1^s score of 0.2302, followed by *PredComX* [24] (0.2215) and *Miss Qiu* [25] (0.2169). The overall performance of the top 25 teams is visualized in Figure 3.

In comparison, the best-performing team in 2024 (also *webmaking*) achieved a much higher F_1^s score of 0.4089, with over 20 teams surpassing the 0.30 mark on the final leaderboard. In 2025, however, no team exceeded an F_1^s of 0.24, reflecting a considerably more challenging evaluation scenario. Several factors likely contributed to the overall lower scores, but we attribute the largest impact to the expanded and more ecologically diverse test set, which significantly increased the need for model generalization.

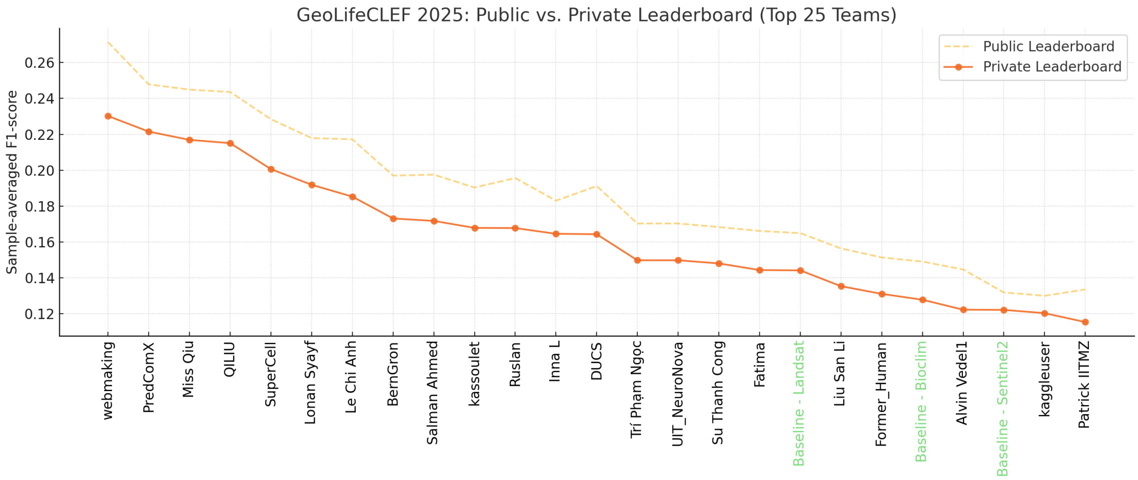

Figure 3: GeoLifeCLEF 2025 results; top 25 teams on public and private leaderboards. Most teams show a drop in performance from public to private scores, likely due to limitations in generalization. The performance gap varies by team, suggesting differing levels of overfitting. Baseline models are highlighted in green.

Unlike the 2024 edition, where the test samples were geographically closer to the training data and many teams relied primarily on PA data, this year's setup required effective use of PO data to succeed, a task that remains difficult due to its inherent biases and lack of negative labels. Overall, while the absolute performance dropped, the technical quality and competitiveness remained high. The challenge successfully pushed participants to develop more generalizable, scalable, and multimodal solutions. Further technical details are available below.

4. Participant's Methods

Out of 41 teams that participated in the GeoLifeCLEF 2025 challenge, 5 submitted working notes reports for peer review. The submitted approaches reflect a diverse set of strategies, including multimodal fusion architectures, rare species handling, spatial post-processing, ensemble learning, and confidence-based filtering. Many top-performing solutions are built upon the baseline models while incorporating additional mechanisms to address class imbalance and spatial shift. Below, we summarize the core techniques used by the participants who submitted their working notes. Full implementation details are available in the respective working notes [23, 24, 25, 26, 27].

Team **webmaking** [23] (Top1) developed a four-component ensemble designed to address the strong class imbalance and spatial shift present in the test data. The approach integrated (i) a multimodal MLP-R + ResNet-18 + EfficientNet-B4 classifier trained on all species, (ii) a rare-species version of the same classifier trained only on infrequent taxa, and (iii) a GeoCLIP model [28] leveraging satellite imagery and metadata. These classifiers were combined with a CatBoost [29] regressor predicting the number of plant species per location. Spatially-aware post-processing using Jaccard-based similarity on a 0.1° grid further improved predictions. The final ensemble, which combined all three classifiers and applied multiple filters, achieved an F_1^s score of 0.2302 and 0.2714 on the private and public leaderboards, respectively.

Team **PredComX** [24] (Top2) introduced a hybrid framework integrating Joint Species Distribution Modeling (JSDM) and deep learning. A ResNet-based deep-SDM extracted features from remote sensing inputs, which were then used to train a Hierarchical Model of Species Communities (HMSC) that modeled interspecies correlations and accounted for study design structure. Their final ensemble included three models: a pure deep-SDM, a pure JSDM, and a combined MLP+HMSC model using features from the deep-SDMs as input to the hierarchical JSDM. The method achieved strong spatial generalization and interpretability, securing second place with an F_1^s score of 0.2215.

Team **Miss Qiu** [25] (Top3): This team proposed Tighnari v2, an improved multimodal framework based on their solution of the previous edition of the challenge [30]. Their approach addresses label noise in PO data through a novel pseudo-label aggregation strategy and mitigates geographic distribution shifts using a mixture-of-experts inference scheme. The model integrates satellite imagery, temporal data, and tabular features via a stackable tri-modal cross-attention module, and employs asymmetric loss [31] to handle class imbalance. Their solution, which achieved 3rd place in this year's edition of the challenge with a F_1^s of 0.218 on the private test set, also outperformed the 2nd-place score from 2024 [20].

Team **Lonan Syayf** [26] (Top6): This participant proposed a multimodal deep learning approach based on three separate Swin-T transformer encoders [32], each specialized for a different input modality: Sentinel-2 imagery, Landsat time series, and bioclimatic rasters. The modality-specific features were projected, concatenated, and passed through an MLP for multi-label species presence prediction [33]. They filtered the label space to plant species with at least 5 PA occurrences and applied a hybrid inference strategy combining a tuned probability threshold (0.18) with a fallback minimum of 14 predictions per site. Their model, trained exclusively on PA data, achieved an F_1^s of 0.192 on the private test set.

Team **BernGron** [27] (Top8): This team combined the PO and PA data in a two-stage deep learning pipeline. They first pre-trained a ResNet18 model [22] on the PO observations to learn general environmental patterns and then fine-tuned it on the PA records for more accurate absence modeling. They tested this strategy across three environmental data modalities: Sentinel-2 imagery, Landsat time series, and bioclimatic variables. Their approach showed that PO-based pretraining improved predictive performance of PA-only baselines, with 7% absolute gains in F_1^s (0.173 on the private test set). They also performed spatial bias analyses using Jensen–Shannon divergence [34] and permutation tests.

5. Discussion and Conclusion

This paper presented an overview and evaluation of the GeoLifeCLEF 2025 challenge, hosted within the LifeCLEF [35, 36] and FGVC12 workshops. Building on previous editions, this year's edition kept its focus on large-scale species distribution modeling using multimodal remote sensing and environmental data. Participants were tasked to predict the presence of plant assemblages at geolocated survey sites using satellite imagery, climatic time series, and tabular environmental descriptors.

GeoLifeCLEF 2025 introduced two significant changes to the task formulation. First, the geographic distribution of the test data was explicitly shifted relative to the training set, emphasizing the need for spatial generalization. Second, the evaluation placed increased weight on detecting rare taxa, many of which had few training observations or were restricted to novel biogeographic regions. These changes increased the modeling complexity compared to previous years, forcing participants to adopt new strategies for spatial extrapolation, confidence calibration, and predicting rare species.

Despite the increased difficulty, participation remained high, with over 40 teams submitting solutions during the competition. A wide variety of modeling pipelines were explored, ranging from multimodal transformer-based architectures and ensemble learning to ecological modeling frameworks such as Joint Species Distribution Models (JSDMs). The main technical outcomes of the challenge are as follows:

- **Generalization across space is a primary limitation for current SDMs.** The introduction of geographically shifted test data revealed substantial performance degradation in models that lacked spatial understanding. Approaches that explicitly accounted for location, through either pre-processing, architecture, or inference, were more robust to bigger geographic shifts.
- **Multimodal data integration improves accuracy.** Consistent with previous editions, models that used more than one environmental modality, e.g., remote sensing imagery, climate time

series, and topographic or land-use variables, outperformed single-source baselines. The challenge confirms the utility of combining complementary data types to model ecological patterns.

- **Ensemble methods provide a practical way to improve performance.** Many top-performing solutions combined multiple specialized models to capture different aspects of the prediction task. Ensembles helped mitigate overfitting, balance predictions for general and rare species, and smooth uncertainty under distributional shifts.
- **Data quality and annotation type still dictate model performance.** Methods trained solely on Presence-Absence data consistently outperformed those relying only on Presence-Only data. Nevertheless, the selective use of PO data, e.g., for pretraining or pseudo-labeling, proved beneficial when handled carefully.
- **Ecologically informed modeling is gaining prominence.** Some participants incorporated principles from community ecology and biogeography, such as species co-occurrence structure and region-specific species pools. These approaches showed promising results and reflect a shift toward more interpretable, hypothesis-driven models in the competition setting.
- **Baselines and community engagement accelerate progress and improve performance.** The continued development of strong open-source baselines and active discourse through the competition platform (Kaggle) enabled participants to iterate quickly, test new hypotheses, and contribute improvements, highlighting the importance of open benchmarking ecosystems in ecological machine learning. Participants extended the baselines by experimenting with alternative architectures, data augmentations, and fusion strategies, demonstrating how shared starting points can accelerate progress and improve overall performance.

Future Directions. While the increasing scale and complexity of the GeoLifeCLEF dataset unlock new research frontiers, it also raises barriers to entry and experimentation. Future editions could explore more modular and accessible task designs, such as regional tracks, taxon-specific subtasks, or single-modality-focused challenges, to maintain broad participation. At the same time, several research directions remain underexplored. First, the development of uncertainty-aware models capable of expressing epistemic uncertainty under geographic or temporal shift would improve both robustness and interpretability. Second, supporting hierarchical taxonomic prediction (e.g., genus-level fallback) could improve performance on rare species. Third, the integration of foundation models trained on environmental data (e.g., SatCLIP [37], BioCLIP [38], GeoCLIP [28]) may offer substantial gains in representation quality. Finally, incorporating ecological priors and spatial constraints, such as species pool filtering or dispersal limitations, could promote more biologically grounded and generalizable model behavior.

6. Declaration on Generative AI

During the preparation of this work, the authors used Grammarly for grammar and spelling checks and ChatGPT for improving clarity and rewording sentences. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

Acknowledgement

The research described in this paper was funded by the European Commission via the MAMBO (http:doi.org/10.3030/101060639) and GUARDEN (http:doi.org/10.3030/101060693) projects, which have received funding from the European Union's Horizon Europe research and innovation program under grant agreements 101060693 and 101060639.

References

[1] C. Botella, A. Joly, P. Bonnet, P. Monestiez, F. Munoz, A deep learning approach to species distribution modelling, Multimedia Tools and Applications for Environmental & Biodiversity Informatics (2018) 169–199.

[2] B. Deneu, A. Joly, P. Bonnet, M. Servajean, F. Munoz, Very high resolution species distribution modeling based on remote sensing imagery: how to capture fine-grained and large-scale vegetation ecology with convolutional neural networks?, Frontiers in plant science 13 (2022) 839279.

[3] J. Estopinan, M. Servajean, P. Bonnet, F. Munoz, A. Joly, Deep species distribution modeling from sentinel-2 image time-series: a global scale analysis on the orchid family, Frontiers in Plant Science 13 (2022) 839327.

[4] E. H. Boakes, P. J. McGowan, R. A. Fuller, D. Chang-qing, N. E. Clark, K. O'Connor, G. M. Mace, Distorted views of biodiversity: spatial and temporal bias in species occurrence data, PLoS biology 8 (2010) e1000385.

[5] N. J. Isaac, M. J. Pocock, Bias and information in biological records, Biological Journal of the Linnean Society 115 (2015) 522–531.

[6] T. Mesaglio, C. T. Callaghan, An overview of the history, current contributions and future outlook of inaturalist in australia, Wildlife Research 48 (2021) 289–303.

[7] C. Garcin, A. Joly, P. Bonnet, J.-C. Lombardo, A. Affouard, M. Chouet, M. Servajean, T. Lorieul, J. Salmon, Pl@ ntnet-300k: a plant image dataset with high label ambiguity and a long-tailed distribution, in: NeurIPS 2021-35th Conference on Neural Information Processing Systems, 2021.

[8] B. Deneu, T. Lorieul, E. Cole, M. Servajean, C. Botella, P. Bonnet, A. Joly, Overview of lifeclef location-based species prediction task 2020 (geolifeclef), CEUR-WS, 2020.

[9] C. Botella, P. Bonnet, A. Joly, Overview of GeoLifeCLEF 2018: location-based species recommendation, in: CLEF task overview 2018, CLEF: Conference and Labs of the Evaluation Forum, Sep. 2018, Avignon, France., 2018.

[10] T. Lorieul, E. Cole, B. Deneu, M. Servajean, A. Joly, Overview of GeoLifeCLEF 2021: Predicting species distribution from 2 million remote sensing images, in: Working Notes of CLEF 2021 - Conference and Labs of the Evaluation Forum, 2021.

[11] T. Lorieul, E. Cole, B. Deneu, M. Servajean, A. Joly, Overview of GeoLifeCLEF 2022: Predicting species presence from multi-modal remote sensing, bioclimatic and pedologic data, in: Working Notes of CLEF 2022 - Conference and Labs of the Evaluation Forum, 2022.

[12] C. Botella, M. Servajean, P. Bonnet, A. Joly, Overview of GeoLifeCLEF 2019: plant species prediction using environment and animal occurrences, CLEF: Conference and Labs of the Evaluation Forum (2019).

[13] C. Botella, B. Deneu, D. Marcos, M. Servajean, T. Larcher, C. Leblanc, J. Estopinan, P. Bonnet, A. Joly, Overview of geolifeclef 2023: Species composition prediction with high spatial resolution at continental scale using remote sensing, in: CLEF 2023 Working Notes-24th Conference and Labs of the Evaluation Forum, volume 3497, 2023, pp. 1954–1971.

[14] A. Joly, H. Goëau, H. Glotin, C. Spampinato, P. Bonnet, W.-P. Vellinga, R. Planque, A. Rauber, B. Fisher, H. Müller, LifeCLEF 2014: Multimedia Life Species Identification Challenges, in: CLEF: Cross-Language Evaluation Forum, number 8685 in Information Access Evaluation. Multilinguality, Multimodality, and Interaction, Springer International Publishing, Sheffield, UK, 2014, pp. 229–249.

[15] A. Joly, H. Goëau, H. Glotin, C. Spampinato, P. Bonnet, W.-P. Vellinga, J.-C. Lombardo, R. Planqué, S. Palazzo, H. Müller, Lifeclef 2017 lab overview: multimedia species identification challenges, in: Experimental IR Meets Multilinguality, Multimodality, and Interaction: 8th International Conference of the CLEF Association, CLEF 2017, Dublin, Ireland, September 11–14, 2017, Proceedings 8, Springer, 2017, pp. 255–274.

[16] A. Joly, H. Goëau, S. Kahl, L. Picek, T. Lorieul, E. Cole, B. Deneu, M. Servajean, A. Durso, H. Glotin, et al., Overview of lifeclef 2022: an evaluation of machine-learning based species identification and species distribution prediction, in: International Conference of the Cross-Language Evaluation Forum for European Languages, Springer, 2022, pp. 257–285.

[17] A. Joly, L. Picek, S. Kahl, H. Goëau, V. Espitalier, C. Botella, B. Deneu, D. Marcos, J. Estopinan, C. Leblanc, T. Larcher, M. Šulc, M. Hrúz, M. Servajean, et al., Overview of LifeCLEF 2024: Challenges on species distribution prediction and identification, in: International Conference of the Cross-Language Evaluation Forum for European Languages, Springer, 2024.
[18] M. Chytrỳ, S. M. Hennekens, B. Jiménez-Alfaro, I. Knollová, J. Dengler, F. Jansen, F. Landucci, J. H. Schaminée, S. Aćić, E. Agrillo, et al., European vegetation archive (eva): an integrated database of european vegetation plots, Applied vegetation science 19 (2016) 173–180.
[19] L. Picek, C. Botella, M. Servajean, C. Leblanc, R. Palard, T. Larcher, B. Deneu, D. Marcos, P. Bonnet, a. joly, Geoplant: Spatial plant species prediction dataset, in: A. Globerson, L. Mackey, D. Belgrave, A. Fan, U. Paquet, J. Tomczak, C. Zhang (Eds.), Advances in Neural Information Processing Systems, volume 37, Curran Associates, Inc., 2024, pp. 126653–126676.
[20] L. Picek, C. Botella, M. Servajean, C. Leblanc, R. Palard, T. Larcher, B. Deneu, D. Marcos, J. Estopinan, P. Bonnet, et al., Overview of geolifeclef 2024: Species composition prediction with high spatial resolution at continental scale using remote sensing, in: CLEF 2024-Working Notes of the 25th Conference and Labs of the Evaluation Forum, 186, CEUR, 2024, pp. 1966–1977.
[21] O. Venter, E. W. Sanderson, A. Magrach, J. R. Allan, J. Beher, K. R. Jones, H. P. Possingham, W. F. Laurance, P. Wood, B. M. Fekete, et al., Global terrestrial human footprint maps for 1993 and 2009, Scientific data 3 (2016) 1–10.
[22] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.
[23] N. Semenova, Addressing class imbalance and spatial shift in geolifeclef 2025, in: Working Notes of CLEF 2025 - Conference and Labs of the Evaluation Forum, 2025.
[24] G. Tikhonov, D. Tikhonov, Synthesizing joint and deep species distribution modeling to enhance spatial prediction of plant communities at continental scale, in: Working Notes of CLEF 2025 - Conference and Labs of the Evaluation Forum, 2025.
[25] H. Liu, Y. Wang, C. Shi, T. Xu, H. Xing, Tighnari v2: Mitigating label noise and distribution shift in multimodal plant distribution prediction via mixture of experts, in: Working Notes of CLEF 2025 - Conference and Labs of the Evaluation Forum, 2025.
[26] A. Syayfetdinov, Swin-t based multimodal networks for geolifeclef 2025, in: Working Notes of CLEF 2025 - Conference and Labs of the Evaluation Forum, 2025.
[27] D. Rawlings, T. Chopard, Enhancing presence-absence identification models using presence-only data, in: Working Notes of CLEF 2025 - Conference and Labs of the Evaluation Forum, 2025.
[28] V. Vivanco Cepeda, G. K. Nayak, M. Shah, Geoclip: Clip-inspired alignment between locations and images for effective worldwide geo-localization, Advances in Neural Information Processing Systems 36 (2023) 8690–8701.
[29] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, A. Gulin, Catboost: unbiased boosting with categorical features, Advances in neural information processing systems 31 (2018).
[30] H. Liu, Z. Tao, P. Jiang, Q. Sun, M. Wan, Tighnari: Multi-modal plant species prediction based on hierarchical cross-attention using graph-based and vision backbone-extracted features, in: Working Notes of CLEF 2024 - Conference and Labs of the Evaluation Forum, 2024.
[31] T. Ridnik, E. Ben-Baruch, N. Zamir, A. Noy, I. Friedman, M. Protter, L. Zelnik-Manor, Asymmetric loss for multi-label classification, in: Proceedings of the IEEE/CVF international conference on computer vision, 2021, pp. 82–91.
[32] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, B. Guo, Swin transformer: Hierarchical vision transformer using shifted windows, in: Proceedings of the IEEE/CVF international conference on computer vision, 2021, pp. 10012–10022.
[33] C. Leblanc, A. Joly, T. Lorieul, M. Servajean, P. Bonnet, Species distribution modeling based on aerial images and environmental features with convolutional neural networks, in: CLEF 2022 Working Notes-23rd Conference and Labs of the Evaluation Forum, volume 3180, 2022, pp. 2123–2150.
[34] J. Lin, Divergence measures based on the shannon entropy, IEEE Transactions on Information theory 37 (2002) 145–151.

[35] L. Picek, S. Kahl, H. Goëau, L. Adam, T. Larcher, C. Leblanc, M. Servajean, K. Janoušková, J. Matas, V. Čermák, K. Papafitsoros, R. Planqué, W.-P. Vellinga, H. Klinck, T. Denton, J. S. Cañas, G. Martellucci, F. Vinatier, P. Bonnet, A. Joly, Overview of lifeclef 2025: Challenges on species presence prediction and identification, and individual animal identification, in: International Conference of the Cross-Language Evaluation Forum for European Languages (CLEF), Springer, 2025.

[36] A. Joly, L. Picek, S. Kahl, H. Goëau, L. Adam, C. Botella, M. Servajean, D. Marcos, C. Leblanc, T. Larcher, et al., Lifeclef 2025 teaser: Challenges on species presence prediction and identification, and individual animal identification, in: European Conference on Information Retrieval, Springer, 2025, pp. 373–381.

[37] K. Klemmer, E. Rolf, C. Robinson, L. Mackey, M. Rußwurm, Satclip: Global, general-purpose location embeddings with satellite imagery, in: Proceedings of the AAAI Conference on Artificial Intelligence, volume 39, 2025, pp. 4347–4355.

[38] S. Stevens, J. Wu, M. J. Thompson, E. G. Campolongo, C. H. Song, D. E. Carlyn, L. Dong, W. M. Dahdul, C. Stewart, T. Berger-Wolf, et al., Bioclip: A vision foundation model for the tree of life, in: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2024, pp. 19412–19424.