Received: 23 January 2024 | Revised: 13 May 2024 | Accepted: 12 July 2024

DOI: 10.1111/avsc.12802

Applied Vegetation Science
IAVS

RESEARCH ARTICLE

A deep-learning framework for enhancing habitat identification based on species composition

César Leblanc[1,2] | Pierre Bonnet[2] | Maximilien Servajean[3] | Milan Chytrý[4] | Svetlana Aćić[5] | Olivier Argagnon[6] | Ariel Bergamini[7] | Idoia Biurrun[8] | Gianmaria Bonari[9] | Juan A. Campos[8] | Andraž Čarni[10,11] | Renata Ćušterevska[12] | Michele De Sanctis[13] | Jürgen Dengler[14,15] | Emmanuel Garbolino[16] | Valentin Golub[17] | Ute Jandt[18,19] | Florian Jansen[20] | Maria Lebedeva[21] | Jonathan Lenoir[22] | Jesper Erenskjold Moeslund[23] | Aaron Pérez-Haase[24] | Remigiusz Pielech[25] | Jozef Šibík[26] | Zvjezdana Stančić[27] | Angela Stanisci[28] | Grzegorz Swacha[29] | Domas Uogintas[30] | Kiril Vassilev[31] | Thomas Wohlgemuth[32] | Alexis Joly[1]

Correspondence
César Leblanc, Inria, LIRMM, Université de Montpellier, CNRS, Montpellier, FR & CIRAD, UMR AMAP, Montpellier, France.
Email: cesar.leblanc@inria.fr

Funding information
The research described in this paper was funded by the European Commission through the GUARDEN (safeGUARDing biodivErsity aNd critical ecosystem services across sectors and scales) project and the MAMBO (Modern Approaches to the Monitoring of BiOdiversity) projects. These projects received funding from the European Union's Horizon Europe research and innovation program under grant agreements 101060693 (start date: 01/11/2022; end date: 31/10/2025) and 101060639 (start date: 01/09/2022; end date: 31/08/2026), respectively. This work had granted access to the High-Performance Computing (HPC) resources of IDRIS (Institut du Développement et des Ressources en Informatique Scientifique) under the allocation 2023-AD011014219 made by GENCI (Grand Equipement National de Calcul Intensif). The authors declare no financial or personal conflicts of interest related to this work. Any ethical considerations

Abstract

Aims: The accurate classification of habitats is essential for effective biodiversity conservation. The goal of this study was to harness the potential of deep learning to advance habitat identification in Europe. We aimed to develop and evaluate models capable of assigning vegetation-plot records to the habitats of the European Nature Information System (EUNIS), a widely used reference framework for European habitat types.

Location: The framework was designed for use in Europe and adjacent areas (e.g., Anatolia, Caucasus).

Methods: We leveraged deep-learning techniques, such as transformers (i.e., models with attention components able to learn contextual relations between categorical and numerical features) that we trained using spatial *k*-fold cross-validation (CV) on vegetation plots sourced from the European Vegetation Archive (EVA), to show that they have great potential for classifying vegetation-plot records. We tested different network architectures, feature encodings, hyperparameter tuning and noise addition strategies to identify the optimal model. We used an independent test set from the National Plant Monitoring Scheme (NPMS) to evaluate its performance and compare its results against the traditional expert systems.

Results: Exploration of the use of deep learning applied to species composition and plot-location criteria for habitat classification led to the development of a framework

For affiliations refer to page 13.

were taken into account during the research process. The views and opinions expressed in this work are those of the authors only and do not necessarily reflect those of the GUARDEN or MAMBO partners, the European Commission, or GENCI.

Co-ordinating Editor: Borja Jiménez-Alfaro

containing a wide range of models. Our selected algorithm, applied to European habitat types, significantly improved habitat classification accuracy, achieving a more than twofold improvement compared to the previous state-of-the-art (SOTA) method on an external data set, clearly outperforming expert systems. The framework is shared and maintained through a GitHub repository.

Conclusions: Our results demonstrate the potential benefits of the adoption of deep learning for improving the accuracy of vegetation classification. They highlight the importance of incorporating advanced technologies into habitat monitoring. These algorithms have shown to be better suited for habitat type prediction than expert systems. They push the accuracy score on a database containing hundreds of thousands of standardized presence/absence European surveys to 88.74%, as assessed by expert judgment. Finally, our results showcase that species dominance is a strong marker of ecosystems and that the exact cover abundance of the flora is not required to train neural networks with predictive performances. The framework we developed can be used by researchers and practitioners to accurately classify habitats.

KEYWORDS

artificial intelligence, biodiversity monitoring, deep learning, European flora, expert system, habitat type identification, phytosociology, species composition, vascular plants, vegetation classification

1 | INTRODUCTION

The term habitat (Hall et al., 1997) encompasses a broad range of definitions (Yapp, 1922). In this study, we adopt the following: “plant and animal communities as the characterizing elements of the biotic environment, together with abiotic factors (soil, climate, water availability and quality, and others), operating together at a particular scale” (Davies & Moss, 1999). The EUNIS habitat classification (Moss, 2008) uses this definition and serves as a comprehensive and hierarchical pan-European system for habitat identification that covers all types of habitats, which are identified by specific codes, names and descriptions. The EUNIS classification system stands nowadays as a widely recognized framework for European habitat types, as it has already played a pivotal role in numerous applications, both research and applied (Evans, 2012). It provides a common language for communication among scientists, policymakers, and other stakeholders. The European Environment Agency (EEA) initiated a (still ongoing) process of revision of the EUNIS habitat classification at the three (and in some cases four) highest levels of its classification hierarchy. This revision led to a more consistent and less ambiguous typology.

Habitat type classification (or identification) is a fundamental process integral to ecology, involving automatically classifying an area based on its environmental characteristics and species composition. It is done by combining observations of species co-occurrence or abundance with environmental estimates to classify vegetation plots across landscapes. Several tools for vegetation classification with different logic and strategy are available, in particular machine-learning algorithms (Hastie et al., 2009) and expert systems (Noble, 1987). The former are tools for induction of the independent knowledge base, whereas the latter emulate the process of expert classification done by humans by using explicitly defined logical formulas. These (numerical) tools can also play a vital role for nature conservation, landscape mapping and land-use planning and can facilitate biodiversity management (Estopinan et al., 2024). They make monitoring of species and habitats easier and more accurate, provide decision support for nature conservation and guidance for nature restoration and development. Thus, it can be particularly valuable in the current context where a significant portion of habitats are at risk of collapsing (at least 32% of European terrestrial habitats and 18% of marine habitats are threatened; Janssen et al., 2016). Therefore, habitat type classification has a crucial role in ecology, and using the EUNIS habitat classification can serve as a key instrument for assessing progress toward the European Union's biodiversity targets.

On the one hand, many expert systems have been published by the global community (Tichý et al., 2019) and have long played a crucial role in protecting and restoring habitats and species. Whether they classify the vegetation of precisely-defined phytosociological units (Marcenò et al., 2018; Novák et al., 2023), the vegetation of entire countries (Chytrý, 2012; Wiser et al., 2018) or even the vegetation or habitats of larger areas (Mucina et al., 2016; Chytrý et al., 2020), these expert systems all follow human decisions. They are usually designed by experts who have extensive knowledge of the characteristics of different habitats and their species composition. These systems thus employ assignment rules (species-based and/or location-based membership conditions) to classify vegetation plots into vegetation or habitat types with formal definitions. However, it is important to note that these definitions can evolve over time, meaning that the structure of the expert systems might need to be modified in order to replace current provisional definitions with improved ones or to use new vegetation-plot records to characterize habitat types. Moreover, the current version of the expert system for automatic classification of European vegetation plots to habitat types of the EUNIS habitat classification (i.e., EUNIS-ESy; Chytrý et al., 2021) contains some definitions that are:

- **strict**, for example, to be correctly assigned to its habitat, a vegetation plot should contain at least *n* species of a given functional species group, or the total cover of a discriminating species group in a vegetation plot should be greater than the total cover of other discriminating species groups in the plot;
- **complex**, for example, to be correctly assigned to its habitat, the total cover of a functional species group in a vegetation plot should be greater than that of another functional group, excluding the species of the former group from the latter group, or the sum of square-rooted percentage covers of the species belonging to a discriminating species group in a vegetation plot should be greater than the sum of square-rooted percentage covers of the species of another discriminating species group;
- and **idiosyncratic**, for example, to be correctly assigned to its habitat, a vegetation plot should belong to a given data set, or a vegetation plot should not be located in a given country.

These intricacies motivate the exploration of alternative approaches, such as the application of deep-learning algorithms, which we delve into in this study.

On the other hand, even though they have shown great potential for modeling species distributions (SDM) (Botella et al., 2018), modern deep-learning techniques have never been applied to classify EUNIS habitats (Joly et al., 2024a), and their application (Černá & Chytrý, 2005) to the classification of habitats at a global scale is a relatively unexplored territory (Joly et al., 2023). Deep-learning techniques are types of machine-learning models that can automatically learn patterns and features from large amounts of data (Botella, Deneu, Gonzalez, et al., 2023) and that are typically designed and trained by data scientists who have expertise in artificial intelligence (AI) and data analysis. As had already been done for species (Deneu et al., 2021), we sought to establish that it was feasible to map the extent of European Union (EU) habitats at (very) high spatial resolution (Deneu et al., 2022). Thus, we used in-situ plant species composition data, information on the location and some environmental features (Leblanc et al., 2022) in a framework with a diverse range of deep-learning models that could be trained for different types of habitats in order to reach an optimal compromise between accuracy and generalization. Habitat type identification has traditionally relied on expert knowledge, a process that is not only time-consuming and costly but also susceptible to subjectivity. Advances in machine learning have opened new opportunities for automating this process using large data sets of environmental and other auxiliary data (Joly et al., 2024b). We built upon these techniques to enable automation and scalability in habitat classification, which forms the cornerstone of our study. AI-powered Habitat Distribution Models (HDMs) should thus be suited to represent how complex ecological niches and spatial dynamics determine the distribution of many habitats in a region. Machine learning could improve predictive performance in HDMs compared to expert systems by better mapping the actual realized distribution of habitat types.

We trained different models on very large volumes of data (by coupling EUNIS types with plant species composition recorded in vegetation plots) to develop, share and maintain a generic, free and open-source deep-learning framework capable of accurately classifying vegetation plots to their habitat types. Several crucial features were introduced into the software package to make it generic and reusable in a wide variety of contexts. We focused our work on five key areas for (i) high modularity (for enhanced flexibility), (ii) new data loaders (to handle both internal and external classification criteria, i.e., respectively species-based and location-based criteria; De Cáceres et al., 2015), (iii) new model's architectures (in particular models based on transformers; Vaswani et al., 2017), (iv) new loss functions (i.e., the penalty for an incorrect classification of a vegetation plot, in particular for species assemblage prediction with an imbalanced top-*k* loss; Garcin et al., 2022) and, (v) a new inference module allowing to compute the top-*k* classification for any user-specified area and plant species composition.

2 | METHODS

2.1 | Data

2.1.1 | EVA: A comprehensive data set for habitat classification

Our data source for training the deep-learning framework was drawn from a subset of the European Vegetation Archive (EVA), a data repository of vegetation-plot observations (i.e., records of plant taxon co-occurrence and cover-abundance at particular sites in plots ranging from 1 m^2 to a few hundred m^2 that have been collected by vegetation scientists) from Europe and adjacent areas. The EVA database (Chytrý et al., 2016), which was accessed on 22 May 2023, is an initiative of the Working Group European Vegetation Survey (EVS). Each of the vegetation plots typically contained estimates of the cover abundance of each species (vascular plants in every vegetation plot, bryophytes and/or lichens in some vegetation plots) alongside various supplementary details and additional sources of information on vegetation structure, location and environmental features. Although the EVA database represents a valuable resource for studying vegetation patterns and dynamics, we considered potential limitations stemming from the representativeness of the data and the possibility of sampling bias (inherent to sets of data assembled from multiple sources and originally collected for various purposes) (Michalcová et al., 2011). The final data set contained a total of 886,260 georeferenced plots (with an average of approximately 20 species per plot), 10,481 different species (see Appendix S2 for the list of all plants species contained in the training data set from EVA) and 228 different habitats (see Appendix S3 for the table listing all habitats from the level three of the EUNIS hierarchy that are present in the EVA training data set), all belonging to one of the eight habitat groups (level one EUNIS habitats) that were the focus of this study, often referred by their 2020 codes:

1. Littoral biogenic habitats (MA2) — 31,533 vegetation plots;
2. Coastal habitats (N) — 37,574 vegetation plots;
3. Wetlands (Q) — 94,100 vegetation plots;
4. Grasslands and lands dominated by forbs, mosses or lichens (R) — 298,816 vegetation plots;
5. Heathlands, scrub and tundra (S) — 67,494 vegetation plots;
6. Forests and other wooded land (T) — 251,474 vegetation plots;
7. Inland habitats with no or little soil and mostly with sparse vegetation (U) — 8018 vegetation plots;
8. Vegetated man-made habitats (V) — 97,251 vegetation plots.

See Appendix S4 for a detailed overview of all the preprocessing steps and to Figure 1 for different visualizations.

2.1.2 | NPMS: An independent data set to evaluate models

To comprehensively assess and compare the transferability of our models and the EUNIS-ESy expert system, we also established an independent and separate test data set whose labels were not generated by the EUNIS expert system or by our algorithms but relied on human annotations. As most of the existing European vegetation-plot databases indexed in the Global Index of Vegetation-Plot Databases (Dengler et al., 2011) (GIVD) and the Global Vegetation Database (Bruelheide et al., 2019) (sPlot) were already included in EVA, obtaining a representative and high-quality independent data set for model validation was challenging. To address this, we selected the National Plant Monitoring Scheme (NPMS) (Walker et al., 2015), which aims to survey plant species across different habitats in the United Kingdom by utilizing data collected by citizens (i.e., expert volunteers who carried out surveys of wildflowers and their associated habitats). This scheme was designed and developed collaboratively by the Botanical Society of Britain & Ireland, UK Centre for Ecology & Hydrology, Plantlife and the Joint Nature Conservation Committee. We specifically chose this data set because it offered an intriguing opportunity to validate the work of numerous European vegetation scientists across generations with a recent citizen science project (Bonnet et al., 2023) that employed a systematic protocol and methodology (e.g., the participants were allocated a 1-km square in which they had to visit five plots in semi-natural habitats twice a year) and encompassed a wide range of vegetation types, providing valuable insights into the potential transferability of our models in a real-world context, beyond expert-driven data sets. It offered an interesting contrast by incorporating data collected through citizen science (Bonnet et al., 2020), thus expanding our understanding of the generalization of the framework beyond traditional scientific data sets. However, this data set is by nature very different from EVA, and there is a significant distribution shift between the two due to the different collection protocols, so we cannot expect the same level of performance. We detail the preprocessing steps to create the test data set in Appendix S4. See Figure 2 for a visual representation of the distribution of the testing data set.

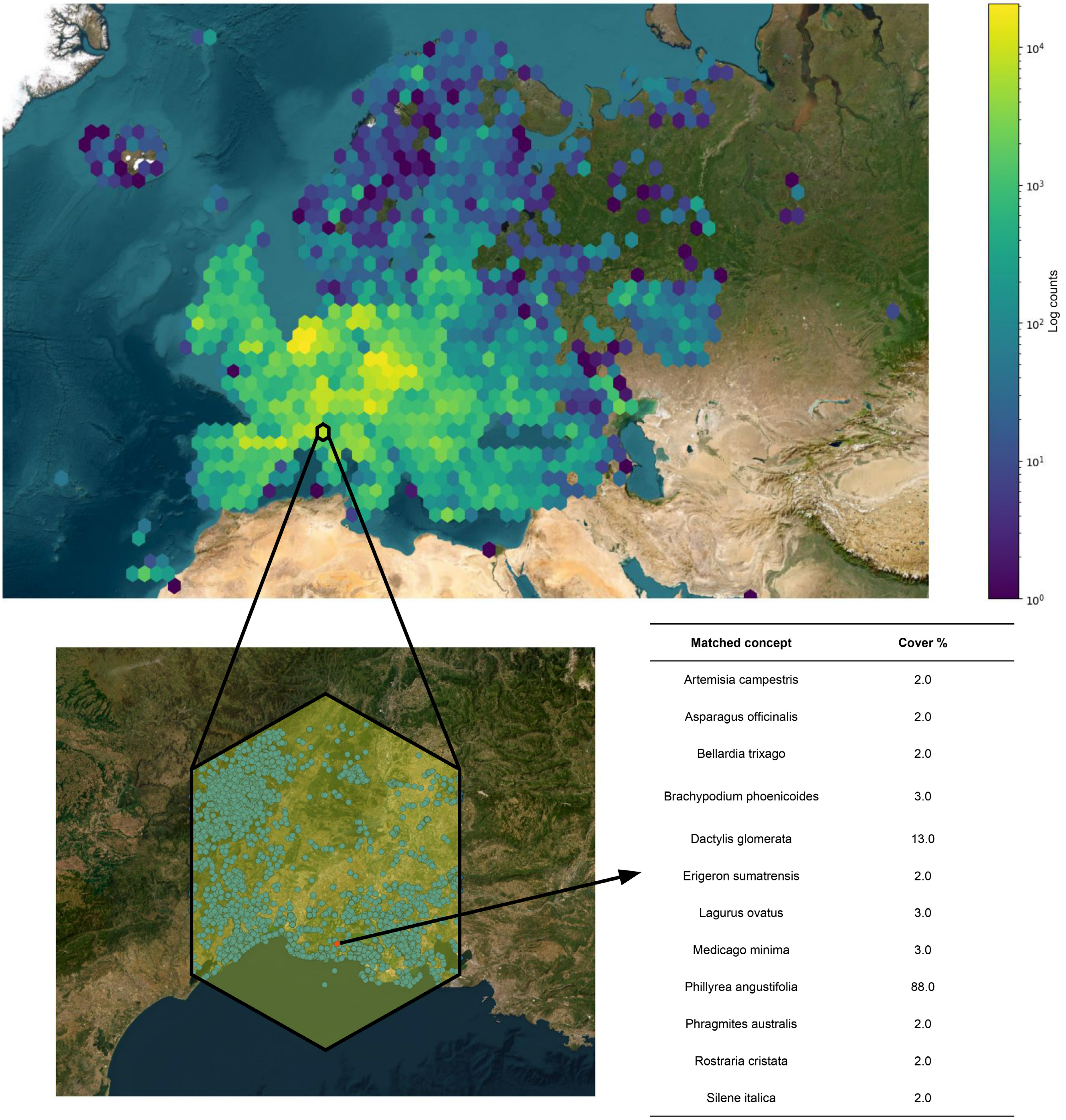
2.2 | Modeling

2.2.1 | Validation: Accounting for the spatial structure of ecological data

The goal of this paper is to use the floristic and environmental information in several locations to train a deep-learning tabular model that can predict the habitat type of given points. To mitigate the influence of spatial autocorrelation and to ensure that our models generalize well beyond the spatial structure of the training data, we split our data set into ten folds according to a spatial block holdout procedure (Roberts et al., 2017). All the vegetation plots were assigned into a grid of 10 km × 10 km cells; all of these cells were then randomly sampled for one of the folds and each fold was used once as an internal validation set while the nine remaining folds formed the training set, allowing us to perform ten-fold cross-validation (CV) (Stone, 1974). The performance measure reported by the ten-fold CV was then the average of the values computed in the loop. This method allowed us to evaluate our approaches in a way that limits the effect of the spatial bias in the data without wasting much of it (which can occur when arbitrarily setting aside a validation set). Importantly, it is worth noting that, regardless of the fold designated for validation in each iteration, every habitat category remained present in the training set formed by the remaining nine folds.

2.2.2 | Models: Using deep neural networks on tabular data for classification

We used the ten-fold CV procedure described above to conduct a rigorous comparative analysis of several machine and deep-learning models. Since there was not an established benchmark for tabular data, we had to work with some of the most used and well-established machine and deep-learning algorithms in competitions, from ensembles of decision trees (Friedman, 2001) to attention-based models (Bahdanau et al., 2014). To ensure fairness and optimize their performances, we meticulously tuned each model's main hyperparameters (for the rest, we kept the default configurations recommended by the corresponding papers) (Feurer & Hutter, 2019). None of the machine and deep-learning models for tabular data described in the existing literature (Borisov et al., 2024) could consistently outperform



Matched concept	Cover %
Artemisia campestris	2.0
Asparagus officinalis	2.0
Bellardia trixago	2.0
Brachypodium phoenicoides	3.0
Dactylis glomerata	13.0
Erigeron sumatrensis	2.0
Lagurus ovatus	3.0
Medicago minima	3.0
Phillyrea angustifolia	88.0
Phragmites australis	2.0
Rostraria cristata	2.0
Silene italica	2.0

FIGURE 1 Hexagonal binning showing the distribution of vegetation plots from the training data set (top), zooming in on a specific bin with the raw spatial distribution of the vegetation plots (bottom), and a vegetation plot (assigned to the habitat type S51, i.e., Mediterranean maquis and arborescent matorral) with the list of co-occurring species.

all the others. To thoroughly evaluate how well our models work, we adopted a variety of approaches and selected neuron-based models (i.e., models that consist of interconnected artificial neurons that learn complex patterns in data through forward and backward propagation), tree-based models (i.e., models that combine multiple base models to improve predictive performance with bagging or boosting) and transformer-based models (i.e., models that enable the capturing of intricate contextual relationships within input data for predictive accuracy). We illustrate each model and the associated training procedure in Appendix S1. Five common models were retained for evaluation:

1. A MultiLayer Perceptron classifier (MLP) (Haykin, 1998), that is, a fully connected class of feedforward artificial neural network. It works by taking input data, passing it through multiple layers of interconnected nodes with weighted connections

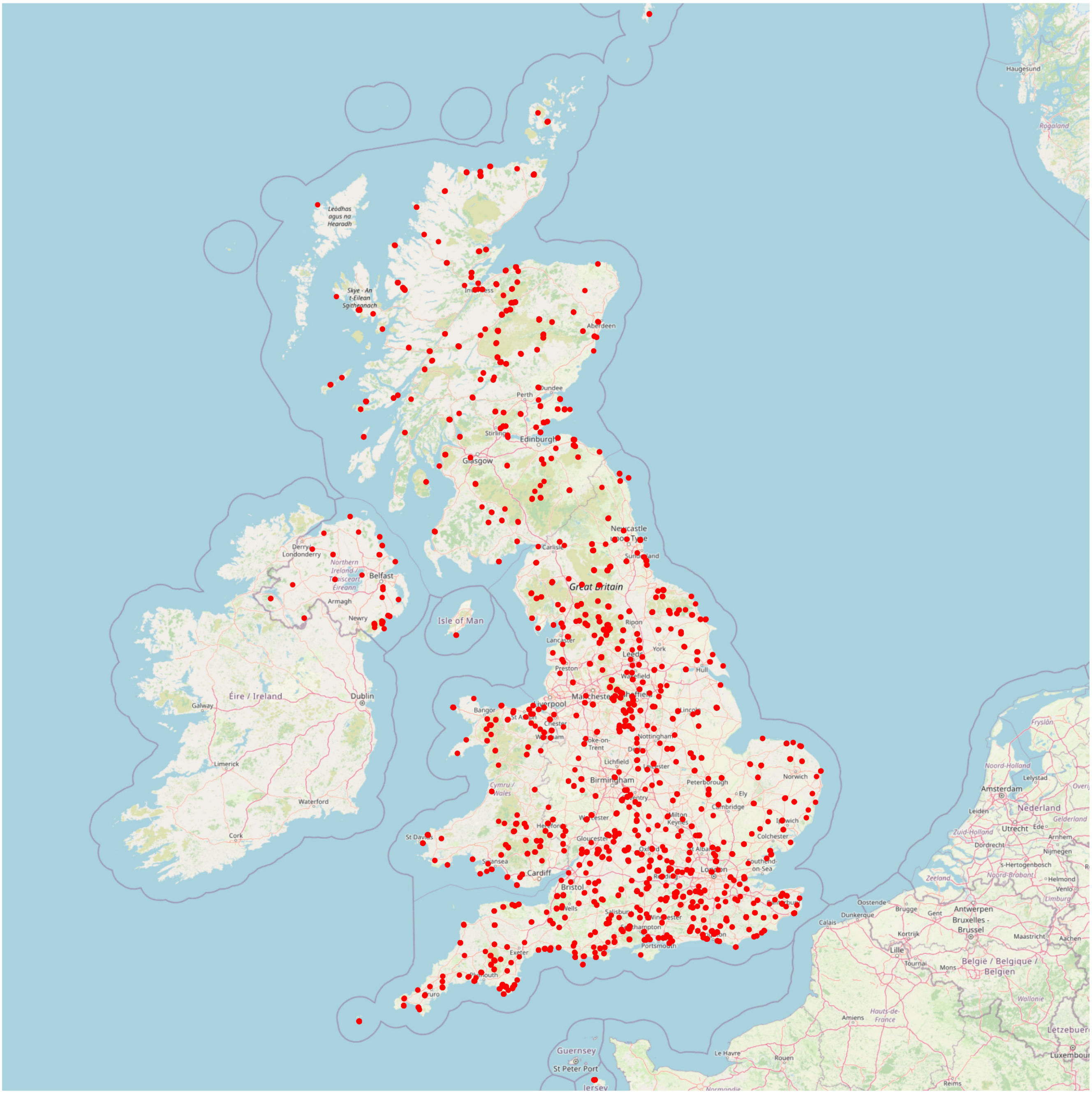


FIGURE 2 Distribution of vegetation plots in the NPMS test set.

and activation functions (Bircanoğlu & Arıca, 2018), and producing output predictions based on the learned patterns in the data.

2. A Random Forest Classifier (RFC) (Ho, 1995), that is, a meta-estimator that fits a number of decision tree classifiers on various subsamples of the data set and uses averaging to improve the predictive accuracy and control overfitting. A single decision tree works by recursively partitioning the input data based on the values of its features to create a tree-like structure, where each internal node represents a feature and each leaf node represents a decision or prediction based on the input data's characteristics.
3. An eXtreme Gradient Boosting classifier (XGB) (Chen & Guestrin, 2016), that is, an optimized distributed gradient-boosting algorithm designed to be highly efficient, flexible and portable. It works by iteratively training and adding decision trees to the ensemble model, each focusing on reducing the residual errors of the previous trees, using a combination of gradient descent optimization (Ruder, 2016), regularization techniques, and hardware-aware optimization to achieve high accuracy and scalability.
4. A TabNet Classifier (TNC) (Arik & Pfister, 2019), that is, a novel high-performance and interpretable canonical deep tabular data learning architecture. It works by selectively attending to the most informative features of the input data and using a sparse

masking technique to allow for efficient and interpretable feature selection, while employing a multistep decision-making process and auxiliary loss functions to enhance its performance and generalization.

5. A Feature Tokenizer+Transformer classifier (FTT) (Gorishniy et al., 2021), that is, a model that transforms all features (categorical and numerical) to embeddings and applies a stack of transformer layers to the embeddings. It works by transforming all features to tokens and running a stack of transformer layers over the tokens, so every transformer layer operates on the feature level of one object.

2.2.3 | Encodings: Mapping current habitat distributions under different constraints

The vegetation plots found within EVA contain comprehensive records of plant species co-occurrences and abundances. All categorical variables (i.e., the country name, the terrestrial ecoregion, the coastline and the location on a coastal dune) are transformed using the simple and widely used one-hot encoding technique (Hancock & Khoshgoftaar, 2020). This is an encoding method in which a particular value of a categorical variable having *n* possible categories would be encoded with a one-dimensional feature vector of length *n* where every component is zero except for the *i*th component, corresponding to the index of the particular category in the set of possible values, which has the value one. All numerical features (i.e., the degrees of latitude and longitude and the altitude of the vegetation plot in meters above sea level) were left untouched. We proposed different data representations (as it is known that this can be vital for the success or failure of models; Bengio et al., 2013) to ensure the framework's applicability to both abundance and presence/absence surveys (Joseph et al., 2006). Three distinct techniques for plant species encoding were employed:

1. The cover-abundance of each species, that is, the natural logarithm of the raw data from EVA. In most cases, it was originally recorded using a cover-abundance scale (Westhoff & van der Maarel, 1978). The scale values were transformed to the arithmetic mid-point percent cover value corresponding to the individual cover-abundance class following the default values in the Turboveg database management program (Hennekens & Schaminée, 2001).
2. The presence/absence of each species, that is, the binarization of the raw data from EVA. Each non-zero entry from the original data was converted to the value one, and every explicit zero was preserved (Scherrer et al., 2020).
3. The reciprocal rank of each species, that is, the inverse of the ordinal ranking of the raw data from EVA. Each species was ranked in descending order of its original cover-abundance value (Brun et al., 2023) (from highest to lowest) and was then associated with the value of the inverse of its position in the ranking.

2.3 | Evaluation

2.3.1 | Fitting: Evaluating modeling algorithms on selected covariates

All details about the models and their optimization are provided in Appendix S1. We evaluated the performance of the expert system on the training set we created. EVA data were classified using the EUNIS-ESy expert system (using its definitions of individual EUNIS habitats based on their species composition and geographic location) but we wanted to see if the vegetation plots would remain classified to the same habitat after interpreting the taxon names with the Global Biodiversity Information Facility (GBIF). We thus kept the same 886,260 vegetation plots, we took the names from the original database and proceeded to standardize them. Furthermore, unlike our experiments for which we kept only vascular plant species and species that were observed at least ten times, we also kept in this case species belonging to other phyla (especially bryophytes and lichens since they were used by the expert system in the definition of some habitats such as S12, i.e., moss and lichen tundra) and rare species (as rare species with occurrences concentrated in a particular habitat could be used as positive indicators of the habitat by the expert system). This process increased the number of species observations to 18,867,936 (instead of the 17,718,306 used to evaluate our models) and the number of different species to 17,885 (instead of the 10,481 used to evaluate our models). Two of the 886,260 vegetation plots had no species left after the species name matching, and as the expert system (unlike our framework) cannot classify vegetation plots solely based on external criteria, we added for both vegetation plots a fake species named “Unknown species” having a percentage cover of 10%.

2.3.2 | Metrics: Computing accuracy to evaluate how well the models are performing

Some of the vegetation plots that were automatically classified by EUNIS-ESy were assigned to two or more level-three EUNIS habitats. In order to deal with that and to evaluate the effectiveness of our classification framework, considering the complexity of the habitat classification task, two key metrics were selected:

1. The top-one micro-average multiclass accuracy, that is, $\frac{1}{N}\sum_{i=1}^{N} 1\left(y_i = \hat{y}_i\right)$ where y is the target value (the classification of EUNIS-ESy) and $\hat{y}$ is the predictions (the classification of the framework). It is the conventional accuracy: the model's prediction must be exactly the expected habitat type. This was the most important metric and played a pivotal role in our evaluation, as it provided crucial insights into the performance of our approaches when we were predicting which habitat was the most likely to be observed at a given location.

2. The top-three accuracy, that is, $\frac{1}{N}\sum_{i=1}^{N} e_i$ where e_i equals 1 if $\forall k \in \{1,2,3\}, \hat{y}_{i,k} = y_i$ and equals 0 otherwise and where y_i is a single ground-truth label (produced by EUNIS-ESy) and $\hat{y}_{i,k}$ are candidate labels (produced by the framework), both associated to a sample *i*. This means that any of the model's three highest probability predictions must match the expected answer. This metric was useful for assessing the performances of our methods on similar habitats (i.e., habitats that have almost identical species composition and environmental features and are thus hard to distinguish from one another) and on scenarios where a vegetation plot was associated with several different habitat labels.

2.3.3 | Noise: Assessing the robustness and generalization of models

To enhance the robustness (Sietsma & Dow, 1991) of our approaches (to mitigate the risk of the phenomenon of overfitting; Dietterich, 1995), we experimented with the incorporation of controlled noise to the input data. We introduced 30% of dropout, that is, when evaluating the performance of the models, we gave each present species a 30% chance of being randomly considered absent in the input data. This deliberate introduction of noise served the vital purpose of reducing the risk that our models would overfit the noise in the data by memorizing various peculiarities of some vegetation plots. Instead, it encouraged the models to identify more general and transferable patterns, thus bolstering their ability to make accurate predictions across diverse ecological contexts. It also helped to imitate the omission of plant species during vegetation sampling (e.g., if some species were small and not easily visible) (Morrison, 2021). After encoding the data and adding (or not) noise, standardization of the features to a mean of observed values of zero and a standard deviation of one was always initiated (these values were estimated from the training data, and then the transformation was consistently applied across all data sets), as it has been shown that such manipulation can be of benefit to some models by improving the numerical stability of the calculations (Kuhn & Johnson, 2013).

3 | RESULTS

3.1 | Selection: Finding the best-performing model

Table 1 contains a comprehensive overview of all the results we obtained (with the models already tuned), showcasing the performance of each model–encoding combination. Among the various configurations tested, the model–encoding combination with the best results is a MLP coupled with features encoded using the reciprocal rank method. This configuration outperformed other models both with and without noise addition to the data and when measuring the performance with the top-one micro-average multiclass accuracy (since it is the best suited metric in our case, as we want to prioritize the most likely habitat for each vegetation plot).

Moreover, to gain insights into the run time (since all the experiments were conducted under the same conditions and some people may have to use the models in the regime of a low-tuning time budget), we plotted the time–performance characteristic for the models in Figure 3. For each meticulously tuned configuration, we reported both the averaged evaluation performance obtained on the ten CV folds (denoting how well the models can generalize to unseen samples) and the results obtained on the test set (using the models trained on the entire EVA data set, without holding out part of the available data). As the encoding and the noise addition did not significantly affect the evaluation time or the inference time, we only show the time of the models used with the reciprocal rank and without noise addition. We can see that all models, except XGB, have similar evaluation and inference times, so there is no universally superior solution in terms of time resources. These two comparisons (Table 1 and Figure 3) allowed us to make some interesting findings, highlighting the nuanced trade-offs between various models and encodings, and emphasizing the importance of selecting the most appropriate approach based on both performance and runtime considerations:

- Models based on decision tree ensembles, such as RFC or XGB, can still outperform some of the deep-learning models (MLP, TNC and FTT) we kept in our experiments, while requiring either a shorter (RFC) or a significantly longer (XGB) amount of time to train.

TABLE 1 Comparison of the top-one (in bold) and top-three (in italics) micro-average multiclass accuracy averaged over the ten cross-validation (CV) folds for every model and encoding, with and without noise addition (best top-one result overall with and without noise addition in green background shading).

	Ten-fold CV			Ten-fold CV with 30% dropout		
Models	Cover abundance	Presence/absence	Reciprocal rank	Cover abundance	Presence/absence	Reciprocal rank
MLP	**88.33**/*97.99*	**76.69**/*95.78*	**88.74**/*98.55*	**72.12**/*86.46*	**65.83**/*88.22*	**73.20**/*89.19*
RFC	**80.31**/*95.72*	**73.44**/*93.74*	**79.39**/*95.41*	**72.56**/*91.88*	**66.32**/*89.90*	**72.62**/*92.20*
XGB	**88.33**/*98.84*	**76.52**/*96.23*	**86.80**/*98.56*	**73.18**/*88.15*	**64.74**/*86.08*	**72.49**/*88.58*
TNC	**79.02**/*91.55*	**68.73**/*87.99*	**80.22**/*92.24*	**65.75**/*81.17*	**60.37**/*82.04*	**67.20**/*82.95*
FTT	**86.62**/*96.88*	**75.09**/*93.78*	**86.98**/*97.18*	**71.18**/*84.83*	**64.76**/*86.50*	**71.68**/*86.21*

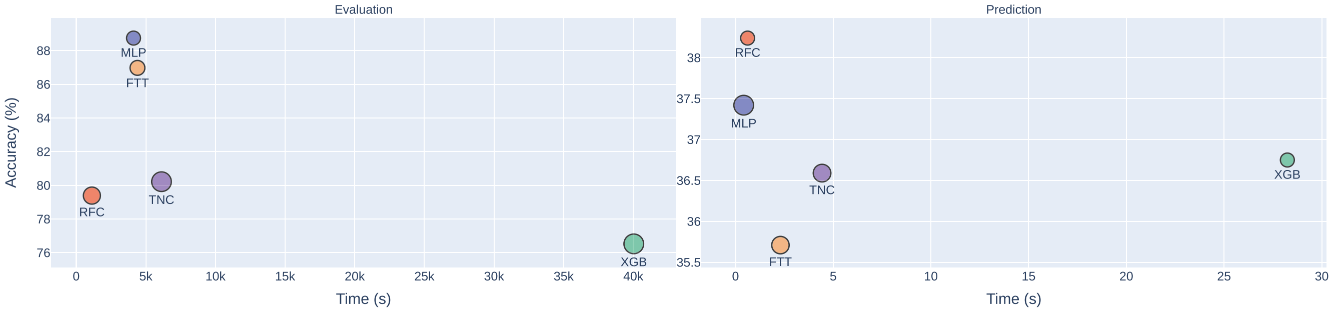


FIGURE 3 Evaluation of time-performance characteristics for selected models on the ten cross-validation folds of the entire EVA training data set of 886,260 vegetation plots (left) and prediction on the NPMS testing data set of 7521 samples (right), with features encoded with the reciprocal rank method (without noise addition). The circle size reflects the top-one micro-average multiclass accuracy standard deviation (left) and the size of the model, that is, the number of trainable parameters for deep-learning algorithms and the number of estimators (i.e., respectively the number of trees in the forest for RFC and the number of gradient-boosted trees for the XGB) for machine-learning algorithms (right).

- Although there has been a clear trend toward transformer-based solutions in recent years, these models, such as TNC and FTT, do not consistently outperform standard neural network architectures, such as MLP.
- The reciprocal rank encoding usually leads to a better top-one performance than the cover abundance (except for tree-based models), despite providing less information about the plant species composition in a given vegetation plot.
- When it comes to top-three performance, there is no ubiquitous encoding technique. However, it seems that there is a trend toward decision tree ensembles (RFC and XGB), as the best models, with and without noise addition, are always tree-based.
- Recent state-of-the-art specialized neural network architectures (e.g., TNC and FTT) and strong traditional machine learning methods (e.g., RFC and XGB) do not provide any benefit over a tuned MLP, which is still more than a simple baseline or a good sanity check (Kadra et al., 2021).

Based on these promising findings, we opted to proceed with the configuration that emerged as the standout performer (i.e., using a MLP classifier with features encoded using the reciprocal rank method and no noise addition) for the subsequent experiments. Indeed, this option was the best trade-off between predictive performance and computational complexity. As the top-three performance of EUNIS-ESy could not be evaluated due to its intrinsic nature, we selected the model with the best top-one accuracy. Moreover, as the EVA data set contained data from all European regions (and not just the UK), we opted for the best-performing model of the ten-fold CV evaluation phase. This strategic choice would be useful for the next phases of our research (i.e., evaluation and interpretability of this configuration and rigorous comparison with the expert system). Having concluded the rigorous process of model selection, which included hyperparameter tuning and the identification of the most effective encoding technique, we proceeded to re-train the chosen model on the entire training data set. This approach allowed us to evaluate the model's performance in a holistic manner (i.e., without partitioning the available data into sets and holding out one of them for evaluation) to compare it to the EUNIS-ESy.

3.2 | Evaluation: Diving into the performance of the best model

Up until now, we employed the micro-average multiclass accuracy to measure the performance of our models. Due to significant class imbalance within the data set (e.g., we had almost 10,000 times more vegetation plots of the R22 habitat than vegetation plots of the R1L habitat in the training set), we aggregated the contributions of all habitats to compute the average metric. However, in some cases, the micro-average may not be the most appropriate metric to evaluate the overall performance of the models. For example, what if we were interested in measuring the performance of the model on each habitat separately, rather than considering the overall performance of the model across all habitats? For such cases, in addition to the introduced metrics, we also computed the macro-average multiclass accuracy metric (still with $k = 1$ and $k = 3$), which is obtained by computing micro-average multiclass accuracy for each class separately and then taking the average over classes. This approach ensured that the habitats with only a few vegetation plots contributed the same as the habitats with thousands of vegetation plots to the assessment of the model's performance. The use of the macro-average multiclass accuracy mitigated the potential issue of smaller classes being overshadowed by larger classes in the overall evaluation of the model's performance.

Before delving into the habitat-specific performance of our model, we conducted further experimentation by training two new MLPs with the reciprocal rank encoding using the same hyperparameters as before, except for one crucial alteration: the reduction applied over labels was replaced by the macro-average. The statistics were calculated for each habitat type (instead of each vegetation plot) and then averaged, but we still used one and three as the numbers of highest probability or logit score predictions considered to find the correct habitat types. There are many more variations between the different folds and a reduction in overall accuracy compared to our previous micro-average results (across all ten CV folds, the model achieved an average multiclass macro-average accuracy of respectively 73.97% and 90.80% for the top-one and top-three metrics, against an average of 88.74% and 98.55% in micro-average accuracy). While our goal was to maintain consistency by employing the same model throughout our experiments, it is important to acknowledge that for habitat-wise performance assessments, it is possible to enhance the results of the MLP model. One promising avenue for improvement is to explore alternative loss functions, for example by switching the currently employed loss function (i.e., the cross-entropy loss; Good, 1952) for the imbalanced top-one and top-three losses, which, after fine-tuning using a grid of parameter values recommended by the authors of the function, outperformed the model's performance under the existing setup.

3.3 | Comparison: Evaluating the performance of hdm-framework and EUNIS-ESy

Of all 886,260 vegetation plots from the data set we used for the expert system, 742,498 were classified to exactly one habitat of level three of one of the eight habitat groups we considered in this study (i.e., MA2, N, Q, R, S, T, U or V). Among the 143,762 other vegetation plots, 11% (i.e., 15,558 vegetation plots) remained unclassified and 4% (i.e., 5748) were classified to more than one habitat. The rest of the vegetation plots (i.e., 122,456 vegetation plots) were classified as habitat groups (i.e., level one habitats), broad habitat types (i.e., level two habitats) or unrevised habitats (i.e., habitats not part of the current EUNIS list). The expert system achieved an accuracy of 85.20%. As the expert system itself was the tool that was used to classify the vegetation plots from EVA, this study shows the lack of robustness to species name standardization of the expert system which clearly overfits the original data.

In addition, we performed a comprehensive performance comparison between our selected model and the expert system on the NPMS test set, presenting the results in Table 2. For this analysis, we used the model trained on the entirety of the EVA data set, without holding out one of the folds for evaluation. We dive deeper into this evaluation exercise and the disparity in performance scores observed between the EVA and the NPMS data sets in Appendix S5.

3.4 | Interpretability: Understanding how the models reason

How the models qualitatively enhance habitat classification is a major question. To answer the increase in model complexity and the resulting lack of transparency, we leveraged different model interpretability methods (e.g., integrated gradients and feature ablation). These are discussed in Appendix S6, in which we dive into the explainability of our models and the ecological interpretability of the results. These state-of-the-art algorithms helped to provide an easy way to understand which features (e.g., which specific plant species) are contributing the most to the model's output. In particular, the implementation of interpretability algorithms can help both researchers and practitioners by facilitating the identification of different plant species that lead the model to assign a vegetation plot to a given habitat type. For example, Figure 4 shows that around 85% of the information about the habitat classification of a vegetation plot is contributed by vascular plant species alone. Additional results include but are not limited to the following:

- The most dominant species inside a vegetation plot are very important in the model's output (e.g., on average, in a vegetation plot containing ten species, over 50% of the total importance is contributed solely by the first two species).
- On average, the model gives more importance to herbaceous species (more than 80%) than to arborescent species (less than 20%), even though this trend is reversed for forests and other wooded land.
- Using solely plant species composition (with neither environmental nor location features) does not decrease the accuracy of the model and it sometimes slightly increases it (e.g., the MLP averages 88.74% with all features and 88.75% with only species composition).

TABLE 2 Performance comparison of hdm-framework and EUNIS-ESy for vegetation-plot classification across various evaluation metrics on the NPMS test set.

	Hdm-framework	EUNIS-ESy
Test accuracy Top-one micro-average multiclass accuracy	37.42%	15.89%
Data requirements Accuracy with neither location nor environmental features	35.39%	15.42%
Representation learning Accuracy with presence/absence data	35.58%	11.63%
Noise robustness Accuracy with 30% dropout	34.50%	13.50%
Calculation speed Time it takes to make predictions	0.42 s	23.91 s

4 | DISCUSSION

4.1 | Main advantages of hdm-framework

We explain in detail the methodology and use of hdm-framework in Appendix S7. For an overview of the primary tasks that can be accomplished using the framework, please refer to Figure 5. Our different experiments have highlighted the remarkable efficacy of AI in classifying vegetation-plot records into their respective EUNIS habitats, marking a significant milestone as the first tool to automate this process across Europe using deep-learning techniques. Notably, our framework not only surpasses the performance of traditional expert systems but also achieves over double the classification accuracy, all while processing data more than 50 times faster than a recently developed electronic expert system. This efficiency carries profound academic and practical implications, benefiting phytosociologists and related fields by potentially expediting research processes and enabling timely conservation initiatives. Furthermore, our work not only underscores the potential of AI within this domain but also points toward a broader paradigm shift in favor of advanced AI solutions. While we acknowledge the need for continued exploration and potential challenges on the horizon, our framework lays a robust foundation for future research and applications in habitat classification. It represents a significant leap forward in the practical utility of the EUNIS habitat classification system.

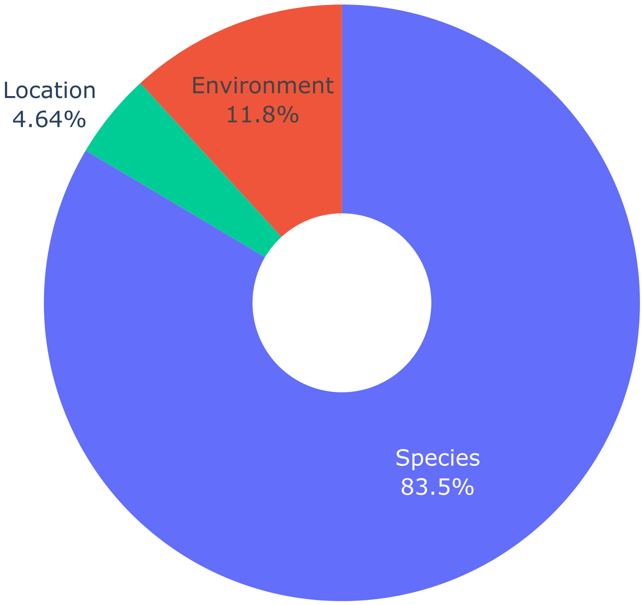


FIGURE 4 Doughnut chart showing the most important group of features of the EVA data set according to the integrated gradients method (applied to the MLP model trained using the reciprocal rank encoding without noise addition). The group "Species" contains the sum of the importance of all species. The group "Environment" contains the sum of the importance of all five environmental variables (i.e., the altitude, the country name, the terrestrial ecoregion, the coastline and the location on a coastal dune). The group "Location" contains the sum of the importance of both longitude and latitude.

EUNIS-ESy, relying on species cover information, encounters limitations when attempting to classify vegetation plots that only record the presence of species without specifying their covers. In contrast, our hdm-framework seamlessly accommodates presence-only data, extending the applicability of such data. Furthermore, traditional expert systems typically assess every vegetation plot within a database, scrutinizing each one to determine if it aligns with one or more predefined habitat definitions specified in their scripts. This process can sometimes lead to vegetation plots remaining unclassified by the expert system. In contrast, the deep-learning models we present in this study were meticulously trained to assign each vegetation plot to (at least) one habitat, which is consistent with the EUNIS habitat classification that was designed to cover all habitat types occurring in Europe.

Hdm-framework is an HDM platform facilitating the use of species occurrence data and environmental features retrieved from multiple sources. Inspired by the existing literature, we proposed several methods that are fast enough to deliver results for thousands of vegetation plots in less than a second. Provided with a set of 195 tunable parameters, hdm-framework has been designed for high customization flexibility, so it can be adapted to anyone's objectives and computing environment. In contrast to the expert system which does not itself extract environmental features, the framework will derive them from the vegetation-plot coordinates using the relevant shapefiles already provided and store the calculated

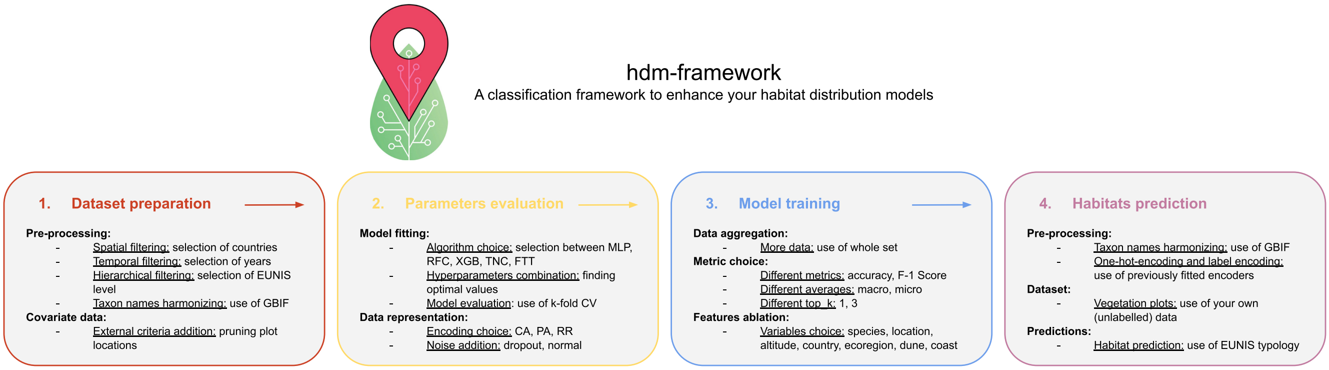

hdm-framework
A classification framework to enhance your habitat distribution models
1. Dataset preparation
Pre-processing:
- Spatial filtering: selection of countries
- Temporal filtering: selection of years
- Hierarchical filtering: selection of EUNIS level
- Taxon names harmonizing: use of GBIF
Covariate data:
- External criteria addition: pruning plot locations
2. Parameters evaluation
Model fitting:
- Algorithm choice: selection between MLP, RFC, XGB, TNC, FTT
- Hyperparameters combination: finding optimal values
- Model evaluation: use of k-fold CV
Data representation:
- Encoding choice: CA, PA, RR
- Noise addition: dropout, normal
3. Model training
Data aggregation:
- More data: use of whole set
Metric choice:
- Different metrics: accuracy, F-1 Score
- Different averages: macro, micro
- Different top_k: 1, 3
Features ablation:
- Variables choice: species, location, altitude, country, ecoregion, dune, coast
4. Habitats prediction
Pre-processing:
- Taxon names harmonizing: use of GBIF
- One-hot-encoding and label encoding: use of previously fitted encoders
Dataset:
- Vegetation plots: use of your own (unlabelled) data
Predictions:
- Habitat prediction: use of EUNIS typology

FIGURE 5 Overview of hdm-framework. The panels display the sequence of tasks performed during each of the four main stages (data set preparation, parameters evaluation, model training and habitats prediction).

values (e.g., location on coastal dunes or in a certain ecoregion) in the header data of the vegetation plots.

4.2 | Potential improvements for practical applications

We discuss the inherent limitations of the training and testing data set in Appendix S4. An essential aspect of our methodology revolves around the standardization of species names using the GBIF Backbone Taxonomy. This step plays a pivotal role in ensuring consistency and facilitating cross-data set comparisons, making it a necessary component of our approach. However, it is important to acknowledge that this process comes with inherent trade-offs, including the loss of valuable information pertaining to species variations and local taxonomic nuances. The harmonization of species names, while promoting uniformity, can inadvertently lead to the amalgamation of distinct taxa or the division of a single taxon into multiple names. Such outcomes have the potential to influence the accuracy of our classification results. Notably, in some instances, phytosociology experts conducting vegetation surveys may have recorded species at a higher taxonomic level, such as specifying the genus (e.g., *Quercus*), without providing precise species designations. This practice presents a challenge during the standardization process, particularly when the GBIF Backbone Taxonomy relies on explicit species information. Consequently, the standardization of higher-level taxonomic names may not always be feasible, potentially impacting the precision of species classification within our framework. Although information recorded at a high taxonomic levels (such as genus) can rarely be informative for habitat classification, it is imperative to recognize and navigate this inherent trade-off between achieving consistency and comparability through species name standardization and the potential loss of finer taxonomic details. This trade-off significantly influences the interpretation and reliability of our classification results, warranting careful consideration in our biodiversity monitoring efforts. Furthermore, the GBIF Application Programming Interface (API) works against data kept in the GBIF Checklist Bank (in partnership with the Catalogue of Life; Bánki et al., 2023) which taxonomically indexes all registered checklist data sets in the GBIF network. It is important to note that this taxonomy store is constantly evolving through updates and takes taxonomic and nomenclatural information from different and new sources, thus potentially resulting in unreproducible results. However, the widespread public deployment of large language models in recent months (Zhao et al., 2023) might offer new opportunities. For example, it could soon be possible to train AI tools on data that have non-standardized nomenclature.

Moreover, the efficacy of our model is intrinsically linked to the taxonomic diversity of vascular plant species present in the training data set (EVA). As our models are trained on this data set, their ability to recognize and classify species is contingent on exposure during training. While in Europe there are more than 20,000 species of vascular plants (Euro+Med, 2006), our framework was trained on a subset comprising 10,481 distinct vascular plants. Consequently, when tasked with classifying plots that contain species not represented in the training set, certain limitations come to the forefront. In instances where our trained models encounter species absent from the training data, it becomes necessary to exclude those unrepresented species as our models would lack familiarity with them. Consequently, this constraint introduces the potential for classification errors, especially in scenarios where a substantial proportion of species within a plot diverges from those within the training set. This limitation is a crucial consideration when applying our framework to novel data sets (Schmidt et al., 2012) or data sets characterized by high species diversity (Botella, Deneu, Marcos, et al., 2023). To enhance the framework's utility and robustness, future endeavors could concentrate on broadening the training set to encompass a more extensive spectrum of species. This expansion could be achieved through various means, including the acquisition of supplementary data sources (Estopinan et al., 2022) or collaboration with domain experts to identify and incorporate missing species (Szymura et al., 2023). Being less cautious during the data curation phase (e.g., by not removing rare species or species with fuzzy or ambiguous names) could also be an option. Exploring strategies to mitigate the impact of species mismatch between training and testing data would be pivotal, further augmenting the framework's versatility and applicability in diverse vegetation classification scenarios.

An essential limitation of our framework pertains to its reliance on predefined habitats for classification. The predictions generated by our models are grounded in the established definitions of EUNIS habitats at the time of model training. In this paper, we focus on eight distinct habitat groups, reflecting the updated EUNIS classification: littoral biogenic habitats, coastal habitats, wetlands, grasslands and lands dominated by forbs, mosses or lichens, heathlands, scrub and tundra, forests and other wooded land, inland habitats with no or little soil and mostly with sparse vegetation, and vegetated man-made habitats. However, it is paramount to recognize that the dynamism of environmental classifications can result in evolving habitat definitions or the emergence of entirely new habitats, driven by agencies such as the EEA. The EUNIS habitat classification itself is currently undergoing a process of revision, and four habitat groups are pending review (inland waters; wetlands; constructed, industrial and other artificial habitats; and complexes). In this respect, leaving some unclassified vegetation plots within the training data could be useful to determine if new habitat types need to be defined. AI techniques could even be used for the definition of those new habitat classes. In addition, climate change (e.g., an increase in temperatures and a decrease in precipitation) and other human influences (e.g., intensification for more productive farming and abandonment of traditional land use) are altering biodiversity, potentially leading to species composition change in some habitats (Blowes et al., 2019). In such cases, our models would necessitate retraining with vegetation plots categorized according to these revised or newly established habitat types. This process can be resource-intensive and potentially environmentally taxing, given the associated energy consumption (Strubell et al., 2020). Therefore, we must acknowledge this limitation and emphasize the importance of periodic model updates to align with any changes in habitat definitions. Furthermore, it underscores the need to consider the ecological footprint of these retraining procedures and explore strategies to optimize their efficiency and sustainability. This may encompass efforts to minimize energy consumption, employ renewable energy sources during the training phase, or investigate eco-friendly training methodologies. By doing so, we can ensure that our framework remains adaptable and environmentally responsible in the face of evolving habitat classifications.

Currently, our framework operates by selecting an integer *k* (by default set to one) and returning the top-*k* habitats with the highest score, a method known as top-*k* classification. Given the complexity of classifying vegetation plots into a substantial number of habitats (a total of 228), relying on a single value for *k* can lead to challenges in precision. To address this issue, we conducted experiments with *k* equal to 3. However, our observations revealed that in cases of high certainty, such as T3B (i.e., *Pinus canariensis* forest, where our MLP model, trained using the reciprocal rank feature encoding method without noise addition, achieved an impressive average top-one micro-average multiclass accuracy of 98.95% across all ten folds), employing *k* larger than 1 resulted in an excessive number of predictions. Conversely, for instances characterized by significant ambiguity, like R1L (i.e., Madeiran oromediterranean siliceous dry grassland, where the same model, trained using the same method, achieved an average accuracy of 0.00% with the same metric and evaluation procedure, although it should be noted that only ten occurrences of this habitat are present in EVA), employing a *k*-value of 3 or less (for example) proved to be overly restrictive. An alternative and promising strategy to address this challenge is the implementation of conformal prediction (Gammerman et al., 2013). This approach dynamically adjusts the number of predicted habitats based on the computed ambiguity for each sample, while still aiming to maintain an average of *k* predictions across all samples, a technique referred to as average-*k* classification (Lorieul et al., 2021). While this approach presents a potential solution for handling ambiguity more effectively, it is important to note that it has not yet been integrated into our framework but represents a promising avenue for future development.

5 | CONCLUSIONS

In summary, the deep-learning framework presented in this paper has demonstrated its remarkable capability to accurately assign vegetation-plot records to their respective EUNIS habitats, as confirmed through rigorous expert evaluation. This framework not only achieves high accuracy and clearly outperforms European expert systems but also ushers in a new era of possibilities. It helps big vegetation data classification and management. The results produced, which are understandable to experts in vegetation classification, highlight the importance of dominant species and the species composition of sites as a whole. The fusion of data sources offers unprecedented flexibility, making it suitable for a wide spectrum of applications across diverse habitat types. For instance, as we consistently assign a substantial number of vegetation plots from various European regions to EUNIS habitat classifications using our framework, it paves the way for precise characterizations of species composition, distribution patterns, and their intricate environmental associations within these habitats. The development of this comprehensive framework represents a significant step toward more efficient, accurate and cost-effective classification of habitat types.

AUTHOR CONTRIBUTIONS

César Leblanc, Pierre Bonnet, Maximilien Servajean and Alexis Joly were involved in the initial idea for the project, helped to define the research questions and objectives and contributed to the overall design of the study; César Leblanc, with contributions from Pierre Bonnet, Maximilien Servajean and Alexis Joly who conducted analyses to compare the performance of the models to the expert system, was responsible for developing the deep-learning framework and ensuring that it was robust, accurate, efficient and well-documented; César Leblanc, with contributions from Pierre Bonnet, Maximilien Servajean and Alexis Joly who ensured that it was of high quality and suitable for the research questions, was responsible for gathering and organizing the data used in the study; César Leblanc was responsible for writing the first draft of the paper and ensuring that it met the standards of the journal; all the other authors were responsible for curating the data delivered from EVA for this study and reviewing and editing the paper; Pierre Bonnet, Maximilien Servajean and Alexis Joly were responsible for overseeing the project as a whole, providing guidance and support to César Leblanc and ensuring that the research was conducted ethically and rigorously; Pierre Bonnet was responsible for securing funding for the project, ensuring that the necessary resources were available to conduct the research. The contributions of each author were integral to the successful execution of this research.

AFFILIATIONS

[1]Inria, LIRMM, Université de Montpellier, CNRS, Montpellier, France
[2]AMAP, Université de Montpellier, CIRAD, CNRS, INRA, IRD, Montpellier, France
[3]LIRMM, AMIS, Université Paul-Valéry - Montpellier 3, CNRS, Montpellier, France
[4]Department of Botany and Zoology, Faculty of Science, Masaryk University, Brno, Czech Republic
[5]Department of Botany, Faculty of Agriculture, University of Belgrade, Belgrade, Serbia
[6]Antenne Languedoc-Roussillon, Conservatoire botanique national méditerranéen, Hyères, France
[7]Swiss Federal Research Institute for Forest, Snow and Landscape Research WSL, Birmensdorf, Switzerland

[8]Department of Plant Biology and Ecology, University of the Basque Country UPV/EHU, Bilbao, Spain
[9]Department of Life Sciences, University of Siena, Siena, Italy
[10]Institute of Biology, Research Center of the Slovenian Academy of Sciences and Art, Ljubljana, Slovenia
[11]School for Viticulture and Enology, University of Nova Gorica, Nova Gorica, Slovenia
[12]Faculty of Natural Sciences and Mathematics, Ss. Cyril and Methodius University, Skopje, Macedonia
[13]Department of Environmental Biology, Sapienza University of Rome, Rome, Italy
[14]Vegetation Ecology Research Group, Institute of Natural Resource Sciences (IUNR), Zurich University of Applied Sciences (ZHAW), Wädenswil, Switzerland
[15]Plant Ecology, Bayreuth Center of Ecology and Environmental Research (BayCEER), University of Bayreuth, Bayreuth, Germany
[16]MINES Paris PSL, ISIGE, Fontainebleau, France
[17]Togliatti, Russia
[18]Geobotany & Botanical Garden, Martin Luther University Halle-Wittenberg, Halle, Germany
[19]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany
[20]Faculty of Agricultural and Environmental Sciences, University of Rostock, Rostock, Germany
[21]Ufa, Russia
[22]UMR CNRS 7058 “Ecologie et Dynamique des Systèmes Anthropisés” (EDYSAN), Université de Picardie Jules Verne, Amiens, France
[23]Department of Ecoscience, Aarhus University, Aarhus, Denmark
[24]Department of Evolutionary Biology, Ecology and Environmental Sciences and Biodiversity Research Institute (IRBio), University of Barcelona, Barcelona, Spain
[25]Institute of Botany, Faculty of Biology, Jagiellonian University, Kraków, Poland
[26]Department of Biodiversity and Ecology, Plant Science and Biodiversity Center, Slovak Academy of Sciences, Bratislava, Slovakia
[27]Faculty of Geotechnical Engineering, University of Zagreb, Varaždin, Croatia
[28]Department of Bioscience and Territory, University of Molise, Termoli, Italy
[29]Botanical Garden, University of Wrocław, Wrocław, Poland
[30]Nature Research Centre, Vilnius, Lithuania
[31]Institute of Biodiversity and Ecosystem Research, Bulgarian Academy of Sciences, Sofia, Bulgaria
[32]Research Unit Forest Dynamics, Swiss Federal Research Institute for Forest, Snow and Landscape Research WSL, Birmensdorf, Switzerland

ACKNOWLEDGEMENTS

The authors extend their gratitude to several individuals and organizations whose contributions were instrumental in the completion of this research. In particular, we thank Tetiana Dziuba for the materials provided and respect her refusal to co-author due to the impossibility of cooperation with Russian authors. We acknowledge the Observatoire Pluridisciplinaire des Alpes-Maritimes (OPAL) infrastructure at Université Côte d'Azur for providing essential resources, including high-performance computing facilities, AI computing resources, data storage, and computational support. We also express our appreciation to the dedicated scientists who collected the original vegetation-plot observations within EVA and the generous database owners and representatives who facilitated access to these valuable data. Special thanks to database administrator Ilona Knollová for her assistance in obtaining the data. The NPMS played a pivotal role in organizing and funding this initiative, and we gratefully acknowledge their support. Our deepest thanks go to the volunteers whose data contributions were indispensable for this study's success. All graphical representations in this paper were created using the Plotly (Inc, 2015) and Matplotlib (Hunter, 2007) plotting libraries. We sincerely appreciate the efforts of all contributors to these open-source libraries, which greatly enhanced the quality of our visualizations. Our analysis also benefited from the FloraVeg.EU website (Chytrý et al., 2024), a valuable resource for European vegetation types, habitats, and plant species information. Furthermore, we acknowledge OpenStreetMap (OSM), an open-data platform licensed under the Open Data Commons Open Database License (ODbL) by the OpenStreetMap Foundation (OSMF) that we used to generate all maps presented in this work. We recognize the OSMF community for their dedicated work in maintaining and curating this invaluable geographic resource, which was instrumental in our study. The collective support and resources provided by these organizations and individuals significantly enriched our research endeavors.

FUNDING INFORMATION

The research described in this paper was funded by the European Commission through the GUARDEN (safeGUARDing biodivErsity aNd critical ecosystem services across sectors and scales) project and the MAMBO (Modern Approaches to the Monitoring of BiOdiversity) projects. These projects received funding from the European Union's Horizon Europe research and innovation program under grant agreements 101060693 (start date: 01/11/2022; end date: 31/10/2025) and 101,060639 (start date: 01/09/2022; end date: 31/08/2026), respectively. This work was granted access to the High-Performance Computing (HPC) resources of IDRIS (Institut du Développement et des Ressources en Informatique Scientifique) under the allocation 2023-AD011014219 made by GENCI (Grand Equipement National de Calcul Intensif). The authors declare no financial or personal conflicts of interest related to this work. Any ethical considerations were taken into account during the research process. The views and opinions expressed in this work are those of the authors only and do not necessarily reflect those of the GUARDEN or MAMBO partners, the European Commission, or GENCI.

DATA AVAILABILITY STATEMENT

This article utilizes data from the European Vegetation Archive (EVA), a comprehensive multicontributor database. The EVA data selection used for this project is stored in the EVA archive at https://doi.org/10.58060/QR4B-G979. While we are unable to publicly share the specific data set used due to third-party restrictions, the vegetation plots we utilized are accessible for research purposes. To replicate our results or conduct further analysis, researchers can submit a proposal to the EVA Coordinating Board to download the data from the archive stored under the above-mentioned Digital Object Identifier (DOI). In contrast, the data set from the National Plant Monitoring Scheme (NPMS) is available under the terms of the Open Government Licence v3 (OGL), which permits unrestricted use and reuse. Interested parties can freely access and utilize the NPMS data set, with conditions as specified by the license. For transparency and reproducibility, the scripts used to generate the analyses presented in this paper, along with the corresponding command lines, are publicly available and can be accessed at https://github.com/cesar-leblanc/hdm-framework/tree/main/Experiments.

ORCID

César Leblanc https://orcid.org/0000-0002-5682-8179
Pierre Bonnet https://orcid.org/0000-0002-2828-4389
Maximilien Servajean https://orcid.org/0000-0002-9426-2583
Milan Chytrý https://orcid.org/0000-0002-8122-3075
Svetlana Aćić https://orcid.org/0000-0001-6553-3797
Olivier Argagnon https://orcid.org/0000-0003-2069-7231
Ariel Bergamini https://orcid.org/0000-0001-8816-1420
Idoia Biurrun https://orcid.org/0000-0002-1454-0433
Gianmaria Bonari https://orcid.org/0000-0002-5574-6067
Juan A. Campos https://orcid.org/0000-0001-5992-2753
Andraž Čarni https://orcid.org/0000-0002-8909-4298
Renata Ćušterevska https://orcid.org/0000-0002-3849-6983
Michele De Sanctis https://orcid.org/0000-0002-7280-6199
Jürgen Dengler https://orcid.org/0000-0003-3221-660X
Emmanuel Garbolino https://orcid.org/0000-0002-4954-6069
Valentin Golub https://orcid.org/0000-0003-3973-6608
Ute Jandt https://orcid.org/0000-0002-3177-3669
Florian Jansen https://orcid.org/0000-0002-0331-5185
Maria Lebedeva https://orcid.org/0000-0002-5020-527X
Jonathan Lenoir https://orcid.org/0000-0003-0638-9582
Jesper Erenskjold Moeslund https://orcid.org/0000-0001-8591-7149
Aaron Pérez-Haase https://orcid.org/0000-0002-5974-7374
Remigiusz Pielech https://orcid.org/0000-0001-8879-3305
Jozef Šibík https://orcid.org/0000-0002-5949-862X
Zvjezdana Stančić https://orcid.org/0000-0002-6124-811X
Angela Stanisci https://orcid.org/0000-0002-5302-0932
Grzegorz Swacha https://orcid.org/0000-0002-6380-2954
Domas Uogintas https://orcid.org/0000-0002-3937-1218
Kiril Vassilev https://orcid.org/0000-0003-4376-5575
Thomas Wohlgemuth https://orcid.org/0000-0002-4623-0894
Alexis Joly https://orcid.org/0000-0002-2161-9940

REFERENCES

Arik, S.O. & Pfister, T. (2019) Tabnet: Attentive interpretable tabular learning. *arXiv preprint, arXiv:1908.07442*.

Bahdanau, D., Cho, K. & Bengio, Y. (2014) Neural machine translation by jointly learning to align and translate. *arXiv preprint, arXiv:1409.0473*.

Bánki, O., Hobern, D., Döring, M., Ower, G., Roskov, Y., Hernandez-Robles, D. et al. (2023) Towards a quality assurance and quality control mechanism for species list building. *Biodiversity Information Science and Standards*, 7, e111665.

Bengio, Y., Courville, A. & Vincent, P. (2013) Representation learning: a review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8), 1798–1828.

Bircanoğlu, C. & Arıca, N. (2018) A comparison of activation functions in artificial neural networks. In *2018 26th Signal Processing and Communications Applications Conference (SIU)*, pages 1–4. IEEE.

Blowes, S.A., Supp, S.R., Antão, L.H., Bates, A., Bruelheide, H., Chase, J.M. et al. (2019) The geography of biodiversity change in marine and terrestrial assemblages. *Science*, 366(6463), 339–345.

Bonnet, P., Affouard, A., Lombardo, J.-C., Chouet, M., Gresse, H., Hequet, V. et al. (2023) Synergizing digital, biological, and participatory sciences for global plant species identification: enabling access to a worldwide identification service. *Biodiversity Information Science and Standards*, 7, e112545.

Bonnet, P., Joly, A., Faton, J.-M., Brown, S., Kimiti, D., Deneu, B. et al. (2020) How citizen scientists contribute to monitor protected areas thanks to automatic plant identification tools. *Ecological Solutions and Evidence*, 1(2), e12023.

Borisov, V., Leemann, T., Sebler, K., Haug, J., Pawelczyk, M. & Kasneci, G. (2024) Deep neural networks and tabular data: a survey. *IEEE Transactions on Neural Networks and Learning Systems*, 35(6), 7499–7519.

Botella, C., Deneu, B., Gonzalez, D.M., Servajean, M., Larcher, T., Leblanc, C. et al. (2023) Overview of geolifeclef 2023: species composition prediction with high spatial resolution at continental scale using remote sensing. Working Notes of CLEF.

Botella, C., Deneu, B., Marcos, D., Servajean, M., Estopinan, J., Larcher, T. et al. (2023) The geolifeclef 2023 dataset to evaluate plant species distribution models at high spatial resolution across Europe. *arXiv preprint, arXiv:2308.05121*.

Botella, C., Joly, A., Bonnet, P., Monestiez, P. & Munoz, F. (2018) A deep learning approach to species distribution modelling. In: Joly, A., Vrochidis, S., Karatzas, K., Karppinen, A. & Bonnet, P. (Eds.) *Multimedia tools and applications for environmental & biodiversity informatics*. Cham: Springer, pp. 169–199.

Bruelheide, H., Dengler, J., Jiménez-Alfaro, B., Purschke, O., Hennekens, S.M., Chytrý, M. et al. (2019) Splot A new tool for global vegetation analyses. *Journal of Vegetation Science*, 30(2), 161–186.

Brun, P., Karger, D.N., Zurell, D., Descombes, P., de Witte, L., de Lutio, R. et al. (2023) Rank-based deep learning from citizen-science data to model plant communities. *bioRxiv preprint*, *bioRxiv*:2023.05.30.542843.

Černá, L. & Chytrý, M. (2005) Supervised classification of plant communities with artificial neural networks. *Journal of Vegetation Science*, 16(4), 407–414.

Chen, T. & Guestrin, C. (2016) Xgboost: a scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York: ACM, pp. 785–794.

Chytrý, M. (2012) Vegetation of The Czech Republic: diversity, ecology, history and dynamics. *Preslia*, 84(3), 427–504.

Chytrý, M., Hennekens, S.M., Jiménez-Alfaro, B., Knollová, I., Dengler, J., Jansen, F. et al. (2016) European Vegetation Aachive (EVA): an integrated database of European vegetation plots. *Applied Vegetation Science*, 19(1), 173–180.

Chytrý, M., Řezníčková, M., Novotný, P., Holubová, D., Preislerová, Z., Attorre, F. et al. (2024) FloraVeg.EU — an online database of European vegetation, habitats and flora. *Applied Vegetation Science*, 27, e12798.

Chytrý, M., Tichý, L., Hennekens, S., Knollová, I., Janssen, J., Rodwell, J. et al. (2021) EUNIS-ESy: Expert system for automatic classification of European vegetation plots to EUNIS habitats.

Chytrý, M., Tichýý, L., Hennekens, S.M., Knollová, I., Janssen, J.A., Rodwell, J.S. et al. (2020) EUNIS Habitat Classification: expert system, characteristic species combinations and distribution maps of European habitats. *Applied Vegetation Science*, 23(4), 648–675.

Davies, C. & Moss, D. (1999) EUNIS habitat classification. Final report to the European Topic Centre on Nature Conservation. Copenhagen: European Environment Agency.

De Cáceres, M., Chytrý, M., Agrillo, E., Attorre, F., Botta-Dukát, Z., Capelo, J. et al. (2015) A comparative framework for broad-scale plot-based vegetation classification. *Applied Vegetation Science*, 18(4), 543–560.

Deneu, B., Joly, A., Bonnet, P., Servajean, M. & Munoz, F. (2022) 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, 839279.

Deneu, B., Servajean, M., Bonnet, P., Botella, C., Munoz, F. & Joly, A. (2021) Convolutional neural networks improve species distribution modelling by capturing the spatial structure of the environment. *PLoS Computational Biology*, 17(4), e1008856.

Dengler, J., Jansen, F., Glöckler, F., Peet, R.K., De Cáceres, M., Chytrý, M. et al. (2011) The Global Index of Vegetation-plot Databases (GIVD): a new resource for vegetation science. *Journal of Vegetation Science*, 22(4), 582–597.

Dietterich, T. (1995) Overfitting and undercomputing in machine learning. *ACM Computing Surveys (CSUR)*, 27(3), 326–327.

Estopinan, J., Bonnet, P., Servajean, M., Munoz, F. & Joly, A. (2024) Modelling species distributions with deep learning to predict plant extinction risk and assess climate change impacts. *arXiv preprint*, *arXiv:2401.05470*.

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

Euro+Med, E. (2006) Euro+Med PlantBbase – the information resource for Euro-Mediterranean plant diversity. Available at: https://www.europlusmed.org

Evans, D. (2012) The EUNIS habitats classification – past, present & future. *Revista de Investigaciones Marinas*, 19(2), 28–29.

Feurer, M. & Hutter, F. (2019) Hyperparameter optimization. In: Hutter, F., Kotthoff, L. & Vanschoren, J. (Eds.) *Automated machine learning: Methods, systems, challenges*. Cham: Springer, pp. 3–33.

Friedman, J.H. (2001) Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, 29, 1189–1232.

Gammerman, A., Vovk, V. & Vapnik, V. (2013) Learning by transduction. *arXiv preprint*, *arXiv:1301.7375*.

Garcin, C., Servajean, M., Joly, A. & Salmon, J. (2022) Stochastic smoothing of the top-k calibrated hinge loss for deep imbalanced classification. In: Proceedings of the 39th International Conference on Machine Learning, PMLR, 162, 7208–7222.

Good, I.J. (1952) Rational decisions. *Journal of the Royal Statistical Society: Series B: Methodological*, 14(1), 107–114.

Gorishniy, Y., Rubachev, I., Khrulkov, V. & Babenko, A. (2021) Revisiting deep learning models for tabular data. *Advances in Neural Information Processing Systems*, 34, 18932–18943.

Hall, L.S., Krausman, P.R. & Morrison, M.L. (1997) The habitat concept and a plea for standard terminology. *Wildlife Society Bulletin*, 25(1), 173–182.

Hancock, J.T. & Khoshgoftaar, T.M. (2020) Survey on categorical data for neural networks. *Journal of Big Data*, 7(1), 1–41.

Hastie, T., Tibshirani, R. & Friedman, J.H. (2009) *The elements of statistical learning: data mining, inference, and prediction*, Vol. 2. New York: springer, pp. 1–758.

Haykin, S. (1998) *Neural networks: a comprehensive foundation*. Upper Saddle River: Prentice Hall PTR.

Hennekens, S.M. & Schaminée, J.H. (2001) TURBOVEG, a comprehensive data base management system for vegetation data. *Journal of Vegetation Science*, 12(4), 589–591.

Ho, T.K. (1995) Random decision forests. In *Proceedings of 3rd International Conference on Document Analysis and Recognition*, volume 1, pp. 278–282. Montreal: IEEE.

Hunter, J.D. (2007) Matplotlib: a 2D graphics environment. *Computing in Science & Engineering*, 9(3), 90–95.

Inc, P.T. (2015) *Collaborative data science*. Montreal: Plotly Technologies Inc.

Janssen, J., Rodwell, J., Criado, M.G., Gubbay, S., Haynes, T., Nieto, A. et al. (2016) *European Red List of Habitats*. Luxembourg: Publications Office of the European Union.

Joly, A., Botella, C., Picek, L., Kahl, S., Goëau, H., Deneu, B. et al. (2023) Overview of LifeCLEF 2023: evaluation of AI models for the identification and prediction of birds, plants, snakes and fungi. In: Arampatzis, A., Kanoulas, E., Tsikrika, T., Vrochidis, S. et al. (Eds.) *Experimental IR Meets Multilinguality, Multimodality, and Interaction*. Cham: Springer.

Joly, A., Picek, L., Kahl, S., Goëau, H., Espitalier, V., Botella, C. et al. (2024a) LifeCLEF 2024 teaser: challenges on species distribution prediction and identification. Lecture Notes in Computer Science Advances in Information Retrieval, 2024.

Joly, A., Picek, L., Kahl, S., Goëau, H., Espitalier, V., Botella, C. et al. (2024b) Overview of lifeclef 2024: challenges on species distribution prediction and identification. In: *International Conference of the Cross-language Evaluation Forum for European Languages*. Cham: Springer.

Joseph, L.N., Field, S.A., Wilcox, C. & Possingham, H.P. (2006) Presence–absence versus abundance data for monitoring threatened species. *Conservation Biology*, 20(6), 1679–1687.

Kadra, A., Lindauer, M., Hutter, F. & Grabocka, J. (2021) Well-tuned simple nets excel on tabular datasets. *Advances in Neural Information Processing Systems*, 34, 23928–23941.

Kuhn, M. & Johnson, K. (2013) *Applied predictive modeling*, Vol. 26. New York: Springer, p. 13.

Leblanc, C., Joly, A., Lorieul, T., Servajean, M. & Bonnet, P. (2022) Species distribution modeling based on aerial images and environmental features with convolutional neural networks. In: *Working notes of CLEF 2022 - conference and labs of the evaluation forum*, pp. 2123–2150.

Lorieul, T., Joly, A. & Shasha, D. (2021) Classification under ambiguity: When is average-k better than top-k? *arXiv preprint*, *arXiv:2112.08851*.

Marcenò, C., Guarino, R., Loidi, J., Herrera, M., Isermann, M., Knollová, I. et al. (2018) Classification of European and Mediterranean coastal dune vegetation. *Applied Vegetation Science*, 21(3), 533–559.

Michalcová, D., Lvončík, S., Chytrý, M. & Hájek, O. (2011) Bias in vegetation databases? A comparison of stratified-random and preferential sampling. *Journal of Vegetation Science*, 22(2), 281–291.

Morrison, L.W. (2021) Nonsampling error in vegetation surveys: understanding error types and recommendations for reducing their occurrence. *Plant Ecology*, 222(5), 577–586.

Moss, D. (2008) *EUNIS habitat classification: a guide for users*. Paris: European Topic Centre on Biological Diversity.

Mucina, L., Bültmann, H., Dierßen, K., Theurillat, J.-P., Raus, T., Čarni, A. et al. (2016) Vegetation of Europe: hierarchical floristic classification system of vascular plant, bryophyte, lichen, and algal communities. *Applied Vegetation Science*, 19(Suppl. 1), 3–264.

Noble, I. (1987) The role of expert systems in vegetation science. *Vegetatio*, 69, 115–121.

Novák, P., Willner, W., Biurrun, I., Gholizadeh, H., Heinken, T., Jandt, U. et al. (2023) Classification of European oak–hornbeam forests and related vegetation types. *Applied Vegetation Science*, 26(1), e12712.

Roberts, D.R., Bahn, V., Ciuti, S., Boyce, M.S., Elith, J., Guillera-Arroita, G. et al. (2017) Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, 40(8), 913–929.

Ruder, S. (2016) An overview of gradient descent optimization algorithms. *arXiv preprint*, *arXiv:1609.04747*.

Scherrer, D., Mod, H.K. & Guisan, A. (2020) How to evaluate community predictions without thresholding? *Methods in Ecology and Evolution*, 11(1), 51–63.

Schmidt, M., Janßen, T., Dressler, S., Hahn, K., Hien, M., Konaté, S. et al. (2012) The West African Vegetation Database. *Biodiversity and Ecology*, 4, 105–110.

Sietsma, J. & Dow, R.J. (1991) Creating artificial neural networks that generalize. *Neural Networks*, 4(1), 67–79.

Stone, M. (1974) Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society: Series B: Methodological*, 36(2), 111–133.

Strubell, E., Ganesh, A. & McCallum, A. (2020) Energy and policy considerations for modern deep learning research. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34, 13693–13696.

Szymura, T.H., Kassa, H., Swacha, G., Szymura, M., Zając, A. & Kącki, Z. (2023) Vegetation databases augment but do not replace species distribution atlases in species richness assessment. *Ecological Indicators*, 154, 110876.

Tichý, L., Chytrý, M. & Landucci, F. (2019) GRIMP: a machine-learning method for improving groups of discriminating species in expert systems for vegetation classification. *Journal of Vegetation Science*, 30(1), 5–17.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N. et al. (2017) Attention is all you need. In: von Luxburg, U., Guyon, I., Bengio, S., Wallach, H. & Fergus, R. (Eds.) *NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems*. Red Hook: Curran Associates Inc., pp. 6000–6010.

Walker, K., Pescott, O., Harris, F., Cheffings, C., New, H., Bunch, N. et al. (2015) Making plants count. *British Wildlife*, 26(4), 243–250.

Westhoff, V. & van der Maarel, E. (1978) The braun-blanquet approach. In: Whittaker, R.H. (Ed.) *Classification of plant communities*. Dordrecht: Springer Netherlands, pp. 287–399.

Wiser, S.K. & De Cáceres, M. (2018) New Zealand's plot-based classification of vegetation. *Phytocoenologia*, 48(2), 153–161.

Yapp, R.H. (1922) The concept of habitat. *Journal of Ecology*, 10(1), 1–17.

Zhao, W.X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y. et al. (2023) A survey of large language models. *arXiv preprint*, *arXiv:2303.18223*.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Appendix S1. Visual overview of the architectures, explanations of the parameters, and model evaluation tables containing the list of tuned hyperparameters with the search spaces and optimal values, the list of fixed hyperparameters with the selected values, and the hardware used, time spent, and result obtained for tuning and optimizing each combination.

Appendix S2. List of all plants species contained in the training data set from EVA.

Appendix S3. Table listing all habitats from level three of the EUNIS hierarchy that are present in the EVA training set or NPMS test set, with their codes, names, conservation statuses based on the European Red List of Habitats and the number of training and testing vegetation plots assigned to each of them.

Appendix S4. Preprocessing steps to create the two data sets used to evaluate habitat distribution models and their limitations.

Appendix S5. Comparison and evaluation of the performance of hdm-framework and EUNIS-ESy.

Appendix S6. Understanding how our models reason using interpretability.

Appendix S7. Guide for the classification framework to enhance habitat distribution models.

How to cite this article: Leblanc, C., Bonnet, P., Servajean, M., Chytrý, M., Aćić, S., Argagnon, O. et al. (2024) A deep-learning framework for enhancing habitat identification based on species composition. *Applied Vegetation Science*, 27, e12802. Available from: https://doi.org/10.1111/avsc.12802