# MaskSDM: Adaptive species distribution modeling through data masking

Robin Zbinden[1], Nina van Tiel[1], Gencer Sumbul[1], Benjamin Kellenberger[2], and Devis Tuia[1]

[1] Ecole Polytechnique Fédérale de Lausanne, Switzerland
[2] University College London, England

**Abstract.** Species distribution models (SDMs) correlate species occurrences with environmental conditions and underpin much of ecological research. A key consideration in developing SDMs is selecting the optimal set of environmental predictor variables, which vary depending on the specific application and species involved. Existing SDMs approaches are limited to a fixed set of predictors defined a priori. This becomes problematic whenever predictors are suboptimal for a particular species or research question to be answered, or when some predictors are unavailable at a given location. To address this, we introduce MaskSDM, a versatile approach that allows end-users to choose relevant variables and gain insights into their contributions to predictions. Our approach employs masked data modeling to learn robust data representations. This allows MaskSDM to effectively handle missing data during both training and inference, addressing a common challenge in real-world geospatial datasets. Evaluations against alternative methods demonstrate that MaskSDM offers improved predictive performance and facilitates valuable analyses of variable contributions.

**Keywords:** Species Distribution Modeling · Deep learning · Masked Data Modeling · Ecology

## 1 Introduction

With the ongoing biodiversity crisis, species distribution models (SDMs) play a key role in monitoring and predicting the distribution of species across various ecosystems. These models, which relate species occurrences to environmental conditions [13], are essential for conservation biology [21]. They are of great interest for understanding current species distributions [15], identifying potential habitats [20], and predicting future changes driven by climate change, habitat destruction, and invasive species [2,31]. Such a wide range of applications, however, requires careful determination of problem-specific characteristics, as ecological environments are diverse and modeling challenges manifold. A critical aspect in this regard is the selection of appropriate environmental predictor variables for the model [1,4,14,29]. For instance, the distribution of many species is impacted by the human footprint [16]; however, when aiming to understand species' ecological niches, human influence is typically excluded from the model. Conventional

SDMs are generally incapable of adapting to such divergent research questions using a single model. Being able to explicitly select input variables of SDMs for each specific task in a generic modeling pipeline presents several challenges, which are outlined below.

*First*, existing multi-species distribution models (which learn the distributions of multiple species with a single model [25]) assume the same set of input variables during training and inference. This prevents the selection of input variables based on the differing needs of different species. We argue that end-users of SDMs should ideally decide which input variables are significant for their specific applications and species. *Second*, the SDMs currently used for ecological explanations are often too simplistic [13,30], limiting their ability to capture complex relationships between the input variables. This hinders accurately providing ecological insights by revealing which input variables contribute to the distribution of species and how they do so [6]. *Third*, existing SDMs require the availability of all the considered input variables for the region of interest. However, some input variables may be missing or known to be noisy or uncertain for some locations during training or inference. This calls for multi-modal solutions able to handle situations with variable (numbers of) inputs.

To address these challenges, we introduce MaskSDM, a method inspired by recent advancements in deep learning, which has demonstrated innovative applications in ecology [35] and SDMs [5,7,41]. MaskSDM enables end-users to select relevant variables at inference and provide insights into their contributions to predictions, while effectively handling missing data. To do so, we leverage masked data modeling, which has proven to be effective in learning robust data representations in natural language processing [9], computer vision [22], and various other data types [28]. Masked data modeling involves randomly hiding parts of the input data, forcing the model to learn and infer the missing information from the remaining predictors. This is typically achieved through a reconstruction objective during a pre-training step, where the model predicts the missing parts [9, 22, 27]. We adopt an approach in which the model directly learns the species distribution in a supervised manner but with missing masked input variables. Our MaskSDM approach offers several key benefits:

1. Flexibility for end-users. MaskSDM allows to select only the variables deemed relevant for specific tasks and species of interest.
2. Analysis of variable contributions. MaskSDM enables a detailed analysis of how different input variables contribute to species predictions.
3. Robustness to missing input variables. MaskSDM handles missing input variables during both training and inference. This ensures that all variables can be included in the model, even if some are unavailable for certain samples.

We evaluate MaskSDM using sPlotOpen [34], a global dataset comprising plant species survey data. Our approach consistently outperforms other SDM methods across various sets of input variables. Furthermore, we demonstrate the effectiveness of MaskSDM through examples that showcase its ability to analyze variable contributions and its adaptability in diverse scenarios, paving the way for the creation of a foundation model for SDMs [3].
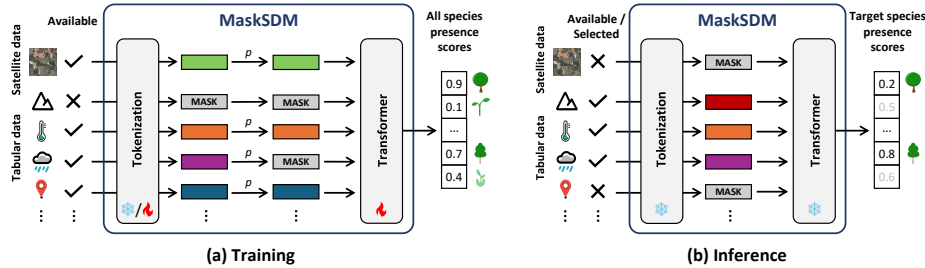
**Fig. 1:** Overview of MaskSDM. (a) During training, our method employs a mask token to indicate missing input variables to the Transformer model. Additionally, this mask token is used to randomly mask each input variable with a probability $p$. (b) During inference, MaskSDM can take any subset of variables as input to predict the presence of species of interest.

## 2 Method

SDMs typically predict the likelihood of observing a species in a given location based on a predefined set of input variables $\mathcal{E} = \{x^1, x^2, \ldots, x^N\}$ representing environmental conditions. These variables are usually assumed to be consistently available and of the same data type [36]. Our approach, called MaskSDM, is capable of learning from multiple types of data inputs, referred to as modalities, while effectively handling missing modalities. This enables MaskSDM to predict the presence of species of interest by exploiting any subset of variables $\mathcal{E}' \subset \mathcal{E}$. To achieve this, we employ masked data modeling to learn species distributions in a supervised manner. The general MaskSDM approach (Fig. 1) is explained in the following, while the specific implementation details are provided in Sec. 3.

**Tokenization** The different modalities are first converted into a standardized format through a process called *tokenization*, which works by projecting inputs to high-dimensional feature vectors (*tokens*) of prescribed size [19, 28]. The functions $f^j$ producing these tokens $t_i^j = f^j(x_i^j)$ are the *tokenizers*, with each modality having a dedicated tokenizer tailored to its specific characteristics. Pre-trained tokenizers, trained on larger datasets, are available for some variables [26, 28]. They can be utilized to produce more informative and general tokens and save computational time. Otherwise, the tokenizers can be trained alongside the rest of the model parameters. While our experiments focus on tabular data and satellite image features as these are commonly used in SDMs [10,17], MaskSDM can be easily extended to incorporate other types of data since each modality has its own tokenizer [28]. Given the diverse distributions of variables in tabular data, we tokenize each input variable individually [18, 19]. Then, to model complex, non-linear interactions between the tokens of the different input variables, all the tokens are given as input to a transformer encoder model [38].

**Masked Data Modeling** During training, MaskSDM employs the masked modeling paradigm to learn robust species distributions. This is achieved by replacing missing input variables with a *mask token* $t_{\mathrm{MASK}}$, signaling to the transformer encoder that a variable is absent. This mask token is learned as part of the model training process, alongside the other parameters of the tokenizers and the transformer encoder. By using mask tokens, we can leverage all available samples during training, even if some values are missing. This is particularly beneficial for real-world datasets, which are often incomplete. For example, we use a dataset containing meta-information (see Sec. 3) that helps explain and disentangle the contributions of different input variables. However, these metadata are not available for all samples. Our approach ensures effective utilization of the dataset despite missing values.

To ensure that the model is robust to any subset of input variables, we randomly mask additional input variables during training, even when they are available [28]. At each iteration, we draw a random number uniformly between 0 and 1, which is the probability of masking each input variable. This stochastic approach forces the model to handle scenarios in which only a limited number of variables are accessible at times, while nearly all variables may be available at others. During inference, MaskSDM offers the capability to replace variables that are missing, or else deemed unsuitable or irrelevant, with the mask token. Moreover, it allows to predict species distributions using different subsets of variables, observing how these choices affect prediction maps and model performance. As a result, users can gain valuable insights into the collective impacts of various input variables on the model's output.

## 3    Experimental Setup

**Dataset** We use the sPlotOpen dataset [34], which comprises 95,104 vegetation plots recorded worldwide. This dataset includes plant species presence and absence data, where the presence of a species is recorded if it is observed in a given plot. To reduce the training time of the model and ensure sufficient data for each species, we focus on predicting the distributions of species with at least 1,000 observations, resulting in a total of 228 species. We use two different methods to divide the data into training, validation, and test sets. The first method involves random splitting of the plots to evaluate the model's *interpolation* capacity. The second method employs spatial blocked cross-validation [32] to assess the model's *extrapolation* capacity (see Appendix A for splits). This approach simulates more realistic scenarios, such as predicting the spread of invasive species [2], ensuring that the model can generalize to new, unseen areas. We collect predictor variables from WorldClim [24] and SoilGrids [23] for each vegetation plot. WorldClim provides essential climate data on temperature and precipitation, widely used as predictors in SDMs [14]. SoilGrids supplies information on soil properties, such as organic carbon content, pH levels, and texture. Additionally, we incorporate human influence data using human footprint maps, encompassing nine variables representing human pressure [39], such as popu-

lation density and nightlights. The longitude and latitude coordinates are also provided to the model, as spatial information has been shown to enhance the performance of SDMs in certain contexts where geographic processes are more influential [11,13]. The sPlotOpen dataset also includes supplementary variables for some plots, such as topographic information (elevation, aspect, slope), and metadata like location uncertainty, surface area, and the coverage and heights of various vegetation layers. This results in a total of 61 tabular data variables, all standardized before input into the model. It is crucial to note that while the WorldClim, SoilGrids, and coordinate variables are consistently available for every plot, other variables are frequently absent (see Table 2 of Sabatini et al. [34]). Finally, image features derived from Sentinel-2 satellite images are incorporated using SatCLIP models [26]. These are available for all plots.

**Model architecture and training** The input variables in the tabular data are tokenized using periodic activation functions [18], followed by a linear layer and ReLU. This approach has demonstrated improved performance in encoding numerical values [18] and is particularly effective for geographic coordinates in capturing features at different scales [33]. For the satellite image features, we utilize the SatCLIP encoder (with 40 Lagrange polynomials) [26], distilled from the ViT16 model [12] trained on satellite images. This encoder serves as the tokenizer, generating tokens that represent characteristics of the satellite images. The transformer encoder follows the architecture used in the FTTransformer [19]. It consists of three identical blocks, each processing tokens of size 192 as both input and output. The number of tokens corresponds to the number of input variables. Each block employs self-attention with 8 heads and a feed-forward network, interleaved with layer normalization and dropout with probability 0.1. Following these blocks, average pooling aggregates the tokens outputted by the final block into a vector of size 192. Subsequently, a linear prediction head with the sigmoid function is applied to generate suitability scores for the 228 species. We optimize the model parameters using stochastic gradient descent with a learning rate of 0.01 and a batch size of 256. To address the class imbalance between presences and absences, we exploit a weighted binary cross-entropy for multi-label classification, with species weights defined according to Zbinden et al. [40]. The model was trained for 300 epochs, with early stopping based on the area under the receiver operating curve (AUC) on the validation set.

## 4   Results

We analyze the effectiveness of MaskSDM compared to common baselines for tabular data (MLP, ResNet, and FTTransformer) [19], where missing values are imputed using the mean of the respective variable [37]. Tab. 1 shows the average AUC across all species on the extrapolation test set using different subsets of input variables. MaskSDM consistently outperforms the baselines, with the performance gap widening as fewer variables are available. Notably, the important

**Table 1:** AUC performance comparison of MaskSDM to different baselines for various subsets of input variables. The number in parentheses indicates the number of input variables in each subset. Note that average temperature is included as part of the WorldClim data. The other subsets are not overlapping.

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Input Variable (#)** | Avg. Temperature (1) | ✓ | ✗ | ✗ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| | WorldClim (19) | ✗ | ✗ | ✗ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| | SoilGrids (8) | ✗ | ✗ | ✗ | ✗ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| | Topographic (3) | ✗ | ✗ | ✗ | ✗ | ✗ | ✓ | ✓ | ✓ | ✓ | ✓ |
| | Location (2) | ✗ | ✓ | ✗ | ✗ | ✗ | ✗ | ✓ | ✓ | ✓ | ✓ |
| | Human footprint (9) | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✓ | ✓ | ✓ |
| | Plot metadata (20) | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✓ | ✓ |
| | Satellite image features | ✗ | ✗ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✓ |
| **Method** | MLP [19] | 69.9 | 75.5 | N/A | 88.1 | 89.0 | 89.7 | 91.1 | 91.2 | 91.5 | N/A |
| | ResNet [19] | 72.5 | 80.7 | N/A | 87.3 | 90.7 | 91.5 | **93.4** | **93.4** | **94.7** | N/A |
| | FTTransformer [19] | 72.2 | 75.3 | 70.2 | 82.1 | 86.0 | 87.3 | 91.8 | 91.9 | 93.7 | 94.3 |
| | MaskSDM (ours) | **80.3** | **88.2** | **88.9** | **91.6** | **92.6** | **93.3** | 93.3 | **93.4** | **94.7** | **94.8** |



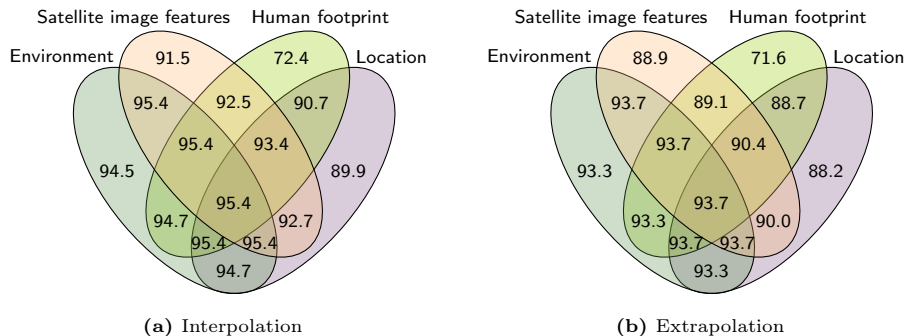**(a)** Interpolation



**(b)** Extrapolation

**Fig. 2:** Comparison of the MaskSDM AUC performance between interpolation and extrapolation testing tasks using different subsets of input variables. The "Environment" subset includes variables from WorldClim, SoilGrids, and topographic data.

difference between MaskSDM and the baseline model with the same architecture (FTTransformer) and training procedure highlights the effectiveness of the masked data modeling approach. This is because the mean imputations in the baselines are treated as genuine data points, while MaskSDM explicitly marks missing variables with a mask token. In addition, the ResNet baseline performs better than the FTTransformer baseline, suggesting that there is potential for improvement in the base implementation of the FTTransformer architecture.

In Fig. 2, we analyze the contributions of different input variables to the performance of the model, comparing the test scenarios of interpolation (Fig. 2a) and extrapolation (Fig. 2b). Unlike traditional ablation studies that require retraining for each combination of input variables, our approach utilizes a single model for all combinations (same for Tab. 1), showcasing the simplicity and efficiency of MaskSDM in analyzing variable contributions. As anticipated, the AUCs are higher for interpolation splits than those for more challenging ex-
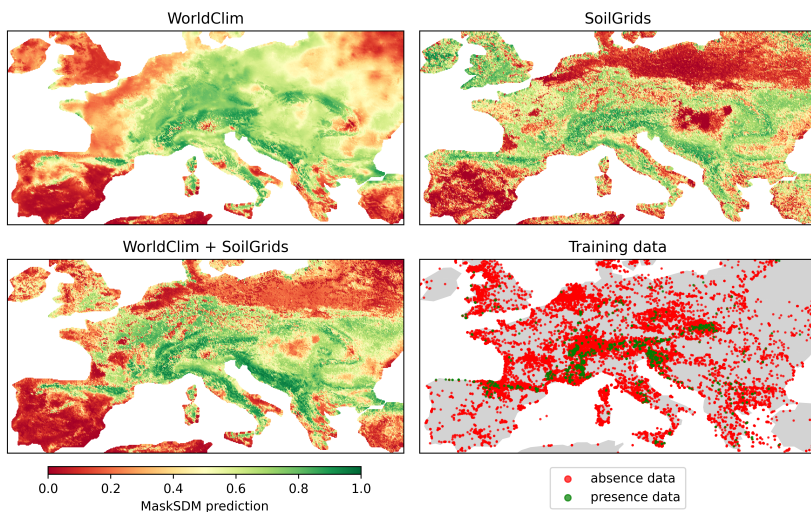
**Fig. 3:** Distribution of *Anthyllis vulneraria* with different subsets of input variables.

trapolation splits. In both cases, environmental variables alone provide strong performance, with modest further improvements seen when adding satellite image features. However, incorporating human footprint variables does not significantly improve performance when combined with other variables. Likewise, location data yields only minimal performance gains for interpolation and no improvements for extrapolation, which aligns with our expectations.

Finally, we illustrate the utility of MaskSDM in visually disentangling the contributions of each variable to the predictions by comparing prediction maps generated using different subsets of input variables, as shown in Fig. 3. We focus on the distribution of *Anthyllis vulneraria*, a medicinal plant native to Europe, whose distribution is significantly influenced by temperature, precipitation, and soil nutrients [8]. The prediction maps confirm the complementary roles of climatic and soil variables, with climatic variables producing coarser predictions compared to the finer-grained predictions of soil variables.

## 5    Conclusion

In this work, we introduce MaskSDM, a flexible approach for handling missing input variables in SDMs that can be tailored to the specific needs of end-users, while facilitating the analysis of variable contributions. By leveraging masked data modeling with transformers, MaskSDM effectively integrates various data sources, even when faced with missing or incomplete modalities. We plan to evaluate MaskSDM with additional data modalities and presence-only datasets, which are prevalent in SDMs. We also aim to conduct in-depth analyses of variable contributions to derive ecological insights, ultimately supporting efforts to address the ongoing biodiversity crisis.

## References

1. Araújo, M.B., Guisan, A.: Five (or so) challenges for species distribution modelling. Journal of biogeography **33**(10), 1677–1688 (2006)
2. Barbet-Massin, M., Rome, Q., Villemant, C., Courchamp, F.: Can species distribution models really predict the expansion of invasive species? PloS one **13**(3), e0193085 (2018)
3. Bommasani, R., Hudson, D.A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M.S., Bohg, J., Bosselut, A., Brunskill, E., et al.: On the opportunities and risks of foundation models. arXiv preprint arXiv:2108.07258 (2021)
4. Bradley, B.A., Olsson, A.D., Wang, O., Dickson, B.G., Pelech, L., Sesnie, S.E., Zachmann, L.J.: Species detection vs. habitat suitability: are we biasing habitat suitability models with remotely sensed data? Ecological Modelling **244**, 57–64 (2012)
5. Brun, P., Karger, D.N., Zurell, D., Descombes, P., de Witte, L.C., de Lutio, R., Wegner, J.D., Zimmermann, N.E.: Multispecies deep learning using citizen science data produces more informative plant community models. Nature Communications **15**(1), 4421 (2024)
6. Bucklin, D.N., Basille, M., Benscoter, A.M., Brandt, L.A., Mazzotti, F.J., Romañach, S.S., Speroterra, C., Watling, J.I.: Comparing species distribution models constructed with different subsets of environmental predictors. Diversity and distributions **21**(1), 23–35 (2015)
7. Cole, E., Van Horn, G., Lange, C., Shepard, A., Leary, P., Perona, P., Loarie, S., Mac Aodha, O.: Spatial Implicit Neural Representations for Global-Scale Species Mapping. In: ICML (2023)
8. Daco, L., Colling, G., Matthies, D.: Altitude and latitude have different effects on population characteristics of the widespread plant anthyllis vulneraria. Oecologia **197**(2), 537–549 (2021)
9. Devlin, J., Chang, M.W., Lee, K., Toutanova, K.: Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805 (2018)
10. Dollinger, J., Brun, P., Sainte Fare Garnot, V., Wegner, J.D.: Sat-sinr: High-resolution species distribution models through satellite imagery. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences **10**, 41–48 (2024)
11. Domisch, S., Friedrichs, M., Hein, T., Borgwardt, F., Wetzig, A., Jähnig, S.C., Langhans, S.D.: Spatially explicit species distribution models: A missed opportunity in conservation planning? Diversity and Distributions **25**(5), 758–769 (2019)
12. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., et al.: An image is worth 16x16 words: Transformers for image recognition at scale. arXiv preprint arXiv:2010.11929 (2020)
13. Elith, J., Leathwick, J.R.: Species distribution models: ecological explanation and prediction across space and time. Annual Review of Ecology, Evolution and Systematics **40**(1), 677–697 (2009)
14. Fourcade, Y., Besnard, A.G., Secondi, J.: Paintings predict the distribution of species, or the challenge of selecting environmental predictors and evaluation statistics. Global Ecology and Biogeography **27**(2), 245–256 (2018)
15. Franklin, J.: Mapping species distributions: spatial inference and prediction. Cambridge University Press (2010)

16. Frans, V.F., Liu, J.: Gaps and opportunities in modelling human influence on species distributions in the anthropocene. Nature Ecology & Evolution pp. 1–13 (2024)
17. Gillespie, L., Ruffley, M., Exposito-Alonso, M.: An image is worth a thousand species: combining neural networks, citizen science, and remote sensing to map biodiversity. bioRxiv pp. 2022–08 (2022)
18. Gorishniy, Y., Rubachev, I., Babenko, A.: On embeddings for numerical features in tabular deep learning. Advances in Neural Information Processing Systems **35**, 24991–25004 (2022)
19. Gorishniy, Y., Rubachev, I., Khrulkov, V., Babenko, A.: Revisiting deep learning models for tabular data. Advances in Neural Information Processing Systems **34**, 18932–18943 (2021)
20. Guisan, A., Thuiller, W., Zimmermann, N.E.: Habitat suitability and distribution models: with applications in R. Cambridge University Press (2017)
21. Guisan, A., Tingley, R., Baumgartner, J.B., Naujokaitis-Lewis, I., Sutcliffe, P.R., Tulloch, A.I., Regan, T.J., Brotons, L., McDonald-Madden, E., Mantyka-Pringle, C., et al.: Predicting species distributions for conservation decisions. Ecology letters **16**(12), 1424–1435 (2013)
22. He, K., Chen, X., Xie, S., Li, Y., Dollár, P., Girshick, R.: Masked autoencoders are scalable vision learners. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 16000–16009 (2022)
23. Hengl, T., Mendes de Jesus, J., Heuvelink, G.B., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M.N., Geng, X., Bauer-Marschallinger, B., et al.: Soilgrids250m: Global gridded soil information based on machine learning. PLoS one **12**(2), e0169748 (2017)
24. Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A.: Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology: A Journal of the Royal Meteorological Society **25**(15), 1965–1978 (2005)
25. Hui, F.K., Warton, D.I., Foster, S.D., Dunstan, P.K.: To mix or not to mix: comparing the predictive performance of mixture models vs. separate species distribution models. Ecology **94**(9), 1913–1919 (2013)
26. Klemmer, K., Rolf, E., Robinson, C., Mackey, L., Rußwurm, M.: Satclip: Global, general-purpose location embeddings with satellite imagery. arXiv preprint arXiv:2311.17179 (2023)
27. Majmundar, K., Goyal, S., Netrapalli, P., Jain, P.: Met: Masked encoding for tabular data. arXiv preprint arXiv:2206.08564 (2022)
28. Mizrahi, D., Bachmann, R., Kar, O., Yeo, T., Gao, M., Dehghan, A., Zamir, A.: 4m: Massively multimodal masked modeling. Advances in Neural Information Processing Systems **36** (2024)
29. Mod, H.K., Scherrer, D., Luoto, M., Guisan, A.: What we use is not what we know: environmental predictors in plant distribution models. Journal of Vegetation Science **27**(6), 1308–1322 (2016)
30. Naas, A.E., Keetz, L.T., Halvorsen, R., Horvath, P., Mienna, I.M., Simensen, T., Bryn, A.: Choice of predictors and complexity for ecosystem distribution models: effects on performance and transferability. Ecography p. e07269
31. Porfirio, L.L., Harris, R.M., Lefroy, E.C., Hugh, S., Gould, S.F., Lee, G., Bindoff, N.L., Mackey, B.: Improving the use of species distribution models in conservation planning and management under climate change. PLoS One **9**(11), e113749 (2014)

32. Roberts, D.R., Bahn, V., Ciuti, S., Boyce, M.S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J.J., Schröder, B., Thuiller, W., et al.: Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography **40**(8), 913–929 (2017)
33. Rußwurm, M., Klemmer, K., Rolf, E., Zbinden, R., Tuia, D.: Geographic location encoding with spherical harmonics and sinusoidal representation networks. In: Proceedings of the International Conference on Learning Representations (ICLR) (2024)
34. Sabatini, F.M., Lenoir, J., Hattab, T., Arnst, E.A., Chytrỳ, M., Dengler, J., De Ruffray, P., Hennekens, S.M., Jandt, U., Jansen, F., et al.: splotopen–an environmentally balanced, open-access, global dataset of vegetation plots. Global Ecology and Biogeography **30**(9), 1740–1764 (2021)
35. Tuia, D., Kellenberger, B., Beery, S., Costelloe, B.R., Zuffi, S., Risse, B., Mathis, A., Mathis, M.W., van Langevelde, F., Burghardt, T., et al.: Perspectives in machine learning for wildlife conservation. Nature communications **13**(1),  792 (2022)
36. Valavi, R., Elith, J., Lahoz-Monfort, J.J., Guillera-Arroita, G.: Modelling species presence-only data with random forests. Ecography **44**(12), 1731–1742 (2021)
37. Van Buuren, S.: Flexible imputation of missing data. CRC press (2018)
38. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I.: Attention is all you need. Advances in neural information processing systems **30** (2017)
39. Venter, O., Sanderson, E.W., Magrach, A., Allan, J.R., Beher, J., Jones, K.R., Possingham, H.P., Laurance, W.F., Wood, P., Fekete, B.M., et al.: Global terrestrial human footprint maps for 1993 and 2009. Scientific data **3**(1), 1–10 (2016)
40. Zbinden, R., Van Tiel, N., Kellenberger, B., Hughes, L., Tuia, D.: On the selection and effectiveness of pseudo-absences for species distribution modeling with deep learning. Ecological Informatics **81**, 102623 (2024)
41. Zbinden, R., Van Tiel, N.M.A., Kellenberger, B.A., Hughes, L., Tuia, D.: Exploring neural networks and their potential for species distribution modeling. In: 11th International Conference on Learning Representations (ICLR) Workshops (2023)