High accuracy is not enough: Interpretability in species distribution model increases with explainable artificial intelligence / interpretable machine learning

Authors
Masahiro Ryo[*1, 2], Boyan Angelov[*,3]
* Equal contribution

Affiliation
[1] Institute of Biology, Freie Universität Berlin, 14195 Berlin, Germany.
[2] Berlin-Brandenburg Institute of Advanced Biodiversity Research (BBIB), 14195 Berlin, Germany.
[3] Association for Computing Machinery (ACM), 1601 Broadway New York, USA.

Keywords
Explainable artificial intelligence; interpretable machine learning; species distribution model; ecological modeling; Local Interpretable Model-agnostic Explanation (LIME)

Abstract
Species distribution modeling (SDM) is a powerful empirical approach for associating the distribution of a species with environmental covariates such as climate factors. However, SDM research still lacks consensus regarding which algorithms are the best for which purposes. We propose that exploring the accuracy-interpretability trade-off using explainable artificial intelligence (xAI; interpretable machine learning) is key to alleviate this issue. xAI is a subfield of AI that aims to enhance the interpretability of complex algorithms. Given an example of SDM combined with xAI (with a reproducible R script), we suggest a promising direction to advance SDM research.

Understanding how and why species are distributed in space is central to biogeography and biological conservation (Pecl et al. 2017, Araújo et al. 2019). Species distribution modeling (SDM) is a powerful empirical approach for associating the distribution of a species with environmental covariates such as climate factors (Elith and Leathwick 2009). Several researchers have put substantial efforts into making SDM more easy-to-build for broad users (cf. >10 software available for SDM as reviewed in Angelov 2019). As a consequence, SDM has matured as a practical ecological modeling tool, and more than 6,000 studies using or mentioning SDM in the past 20 years (Araújo et al. 2019).

While various statistical methods can be used, SDM research still lacks consensus regarding which algorithms are the best for which purposes (Araújo et al. 2019). A key is to find a balance between accuracy and interpretability, given a study objective (Merow et al. 2014). Accuracy and interpretability are typically in a trade-off (Breiman 2001a). To find an optimal balance requires the exploration of both sides along the trade-off gradient: excessively simple algorithms (e.g., linear model) and excessively complex ones (e.g., machine learning and ensemble modeling approach that employs several algorithms together (Araújo and New 2007)).

We argue that it is, therefore, time to reconsider how to assess the 'goodness' of SDM. The performance of an SDM is usually assessed in terms of accuracy. The theoretical rationale of doing so is the bias-variance trade-off. Under this rule, the most complex algorithm is most likely to be the best choice. However, as mentioned above, now SDM research also challenges the accuracy-interpretability trade-off. We think that interpretability is still largely understudied but necessary for bringing more trust in SDM research. Even if an SDM of a target species achieves high accuracy, the user, such as the local ecological manager and decision-maker, may not trust the model when it is difficult to assess whether a model prediction makes sense in the light of the ecology of the target species. Collectively, a truly good SDM should be sufficiently accurate and also sufficiently easy to interpret (Phillips et al. 2004, Austin 2007, Merow et al. 2014). Moreover, we imagine that increasing interpretability may also enhance the realism and generality of SDMs. These are other important 'goodness' features (Araújo et al. 2019).

A complex algorithm is attractive for capturing the complex nature and context-dependency of a dynamic system. But at the same time, it is unattractive because of low interpretability. This dilemma is a common problem in several other disciplines as well as ecology. Any important decision making requires a certain convincing rationale for a prediction (e.g., financial risk assessment, medical diagnosis, and criminal justice, to name a few). The strong need for rationalizing predictions made with artificial intelligence (AI) has led to the rapid emergence of explainable AI (xAI), an AI subfield that aims to enhance the interpretability of complex algorithms. xAI is sometimes interchangeably used as interpretable machine learning (Murdoch et al. 2019). This article first introduces xAI to SDM research.

Among others, a groundbreaking idea was suggested by Ribeiro et al. (2016): Local Interpretable Model-agnostic Explanation (LIME). The eye-pening idea of LIME is to increase the interpretability of an excessively complex model without losing accuracy by reducing the 'unnecessary' degree of the model complexity. A key assumption of LIME, which we believe also applies in ecology, is that the necessary degree of model complexity is scale-dependent. A complex algorithm is suited when modeling a species distribution at the biogeographical scale, but a simpler algorithm is often accurate enough at much smaller scales (e.g. landscape, the scale where the actual management happens). With LIME one can still fit a complex algorithm, but when it comes to the interpretation at a local scale, the algorithm is approximated with a simple model that has a sufficient accuracy at the local scale. The idea of cross-scale model exchange is quite unique and practical.

This innovation additionally offers a novel approach for model validation: site-level assessment (see an example in Box 1). Most complex algorithms cannot provide the reasons for each prediction as an output, and therefore, there is no way to assess how much a prediction at a specific site makes sense ecologically. Imagine that we fit a random forest model to a focal species, and the model predicts its presence at site X. But why does the model think so? Because of the temperature, water availability, or other reasons?—We cannot know. But, xAI tools now make it possible to answer such a question. We can investigate how well the complex model of a target species can offer ecologically convincing reasons on prediction outcomes at some randomly chosen local sites. Site-level model validation in the light of ecology of the species is critically important because (i) even if a prediction seems correct, it can be made for utterly nonsense reasons; (ii) locally important environmental features do not necessarily correspond to globally important environmental features (e.g. microclimate and context-dependence); and (iii) an SDM with a high accuracy score does not necessarily mean highly accurate at the local scale.

We mainly introduced LIME as an example, but more xAI techniques are increasingly available to enhance interpretability and thus trust in complex algorithms: see e.g., Molnar (2019) and Murdoch et al. (2019). xAI has already brought promising successes to other scientific disciplines (e.g. computer vision and medical science). We hope that this article helps make SDM research even more successful with the help of xAI to explore a balance between accuracy and interpretability.

Box 1: Explaining the distribution of African elephant with xAI

Since we hope many ecologists investigate the potential of xAI, we demonstrate how to apply the LIME approach for SDM in R (R Core Team 2017). In this exercise, we make a SDM of *Loxodonta africana* (African elephant). Note that our intention here is neither to advance the ecological knowledge of the species nor to model the species rigorously (e.g. we did not consider spatial autocorrelation). Rather, we aim to show how to implement the technique (the R script is available at https://github.com/boyanangelov/interpretable_sdm). We applied the Random Forest algorithm (Breiman 2001b) for modeling the distribution of *L. africana* (data from GBIF) with bioclimatic covariates (data from WorldClim). For data acquisition and processing we used `sdmbench` package (Angelov 2018), and for model training `mlr` package (Bischl et al. 2016), and for model explanation, `lime` package (Pedersen and Benesty 2019).

Fig. S1 highlights the role of xAI in investigating how the SDM predicts at some local sites. Conventionally, model assessment is based mostly on the visualization of its habitat suitability (upper-left panel), accuracy score and variable importance ranking (lower-left panel). In this example, we interpret that the model is very accurate (AUC = 0.984), and that key covariates are the precipitation amount of the wettest quarter and the temperature of the coldest and driest quarter. This interpretation is important for biogeographical understanding, but it does not help us assess the model reliability at the local scale, where actual conservation happens.

xAI can solve the issue. With LIME we show site-level model validation at three randomly taken sites (right panel). At site A, the model predicted a high presence probability (0.95), and now we ask—*why does the model predict so*? We see that the prediction is supported by all top 5 environmental conditions at the site. At site B, the model also predicts presence (prob.=0.97), but the reasons supporting the prediction differs largely from those of site A. Having different key conditions makes sense, since these sites are so far away (approx. 2,500 km). At site C, the model predicts absence (presence prob. =0.34). Yet, some factors are against the prediction. Hence, at site C, a more careful investigation is needed with a specialist of the species around the site.

As demonstrated, such individual LIME explanations for local sites can help better check spatial variations in covariate importance, which in turn, can contribute to more reasonable conservation decisions, ecological understanding, and higher trust in the model at the local scale.

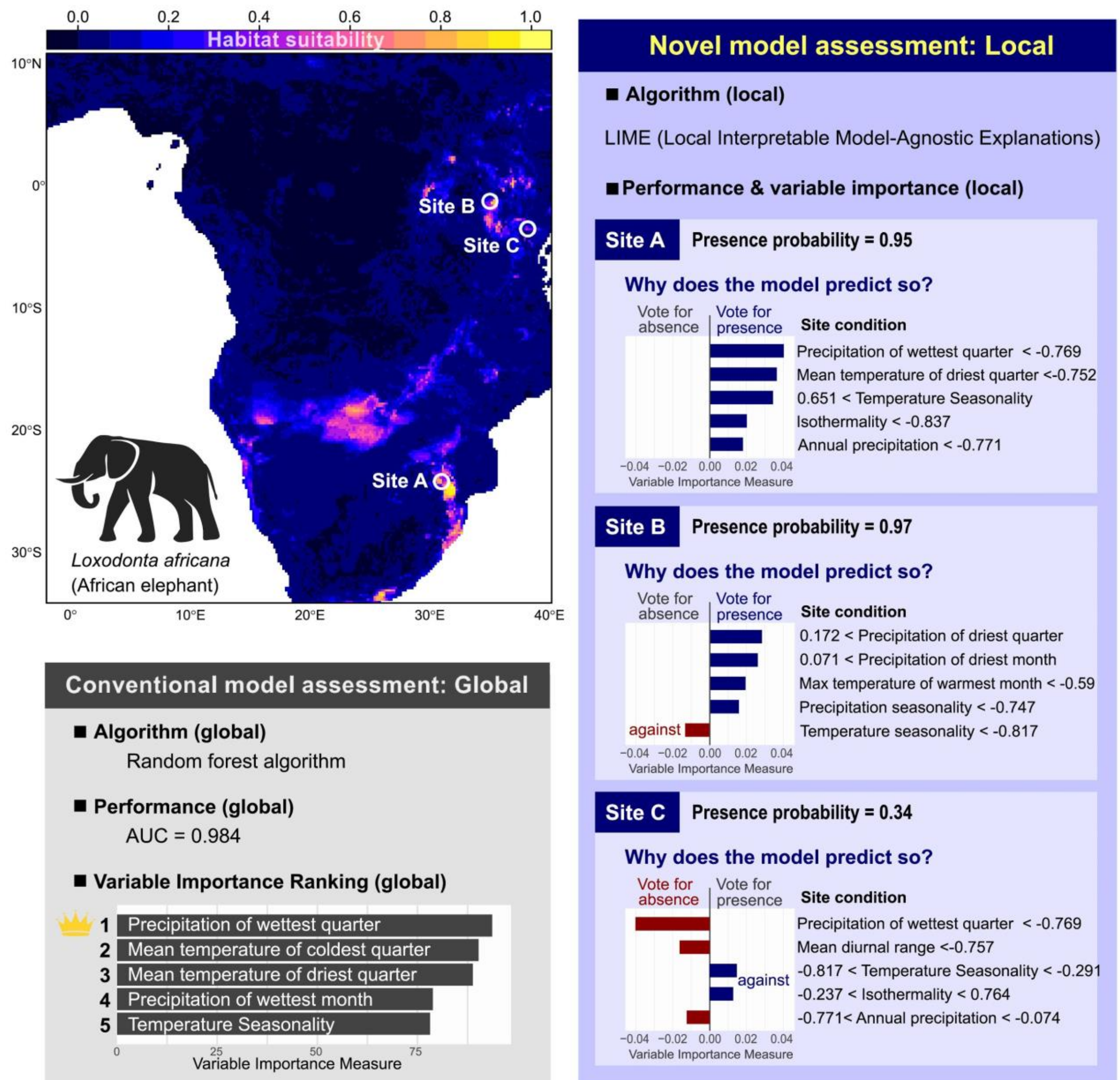

Figure S1. Interpreting the species distribution model of *L. africana* based on habitat suitability mapping and model assessment at both biogeographically and conservationally-relevant scales (global and local). Model interpretation at the local scale applies LIME (local interpretable model-agnostic explanations), an algorithm of explainable artificial intelligence (xAI).

Acknowledgements

MR acknowledges funding by the Grant-in-Aid for JSPS Overseas Research Fellowships.

Author contributions

MR and BA developed the idea, BA wrote the R script and MR contributed for improvement. MR wrote the first draft, and both authors contributed to the writing of the paper.

References

Angelov, B. 2018. sdmbench: R package for benchmarking species distribution models. - J. Open Source Softw. 3: 847.

Angelov, B. 2019. Review of species distribution modeling open-source software.: 5.

Araújo, M. B. and New, M. 2007. Ensemble forecasting of species distributions. - Trends Ecol. Evol. 22: 42–47.

Araújo, M. B. et al. 2019. Standards for distribution models in biodiversity assessments. - Sci. Adv. 5: eaat4858.

Austin, M. 2007. Species distribution models and ecological theory: A critical assessment and some possible new approaches. - Ecol. Model. 200: 1–19.

Bischl, B. et al. 2016. mlr: Machine Learning in R. - J. Mach. Learn. Res. 17: 1–5.

Breiman, L. 2001a. Statistical Modeling: The Two Cultures. - Stat. Sci. 16: 199–215.

Breiman, L. 2001b. Random Forests. - Mach. Lang. 45: 5–32.

Elith, J. and Leathwick, J. R. 2009. Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. - Annu. Rev. Ecol. Evol. Syst. 40: 677–697.

Merow, C. et al. 2014. What do we gain from simplicity versus complexity in species distribution models? - Ecography 37: 1267–1281.

Molnar, C. 2019. Interpretable Machine Learning.

Murdoch, W. J. et al. 2019. Definitions, methods, and applications in interpretable machine learning. - Proc. Natl. Acad. Sci. 116: 22071–22080.

Pecl, G. T. et al. 2017. Biodiversity redistribution under climate change: Impacts on ecosystems and human well-being. - Science in press.

Pedersen, T. L. and Benesty, M. 2019. lime: Local Interpretable Model-Agnostic Explanations.

Phillips, S. J. et al. 2004. A maximum entropy approach to species distribution modeling. - Proc. Twenty-First Int. Conf. Mach. Learn.: 83.

R Core Team 2017. R: A Language and Environment for Statistical Computing. Vienna, Austria. Retrieved from https://www.R-project.org/.

Ribeiro, M. T. et al. 2016. “Why Should I Trust You?”: Explaining the Predictions of Any Classifier. - ArXiv160204938 Cs Stat in press.