Environmental Computational Science and Earth Observation Laboratory

Project subject

From global to local: Modelling the distribution of butterfly species in Switzerland and Belize

Context

Understanding the interaction between climate and the diversity of life is essential to inform and optimize conservation strategies that aim to slow and potentially reverse the pervasive decline in biodiversity[1], as well as the associated destructive consequences for human wellbeing. To this end, Species Distribution Models (SDMs) are key mathematical tools to map the current geographic distributions of species and predicting their future shifts under climate change. These models work by correlating observations of species occurrence with recorded environmental variables[2,3]. Deep learning has emerged as a promising tool for SDMs, leveraging the vast and growing volumes of data generated by citizen science[4,5] and modelling multiple species at the same time[6].

On the one hand, SDMs need to be trained for an adequate spatial extent that covers the entirety of the species range to accurately capture species–environment relationships. On the other hand, biodiversity conservation decisions and policies are mostly taken at a smaller scale, such as regional or local levels, with SDM studies specifically commissioned for limited spatial domains. This mismatch could lead to a spatially truncated representation of the range of the species which implies a lack of accuracy, a limited model transferability and extrapolation under different environmental conditions. Recently, a spatially-nested hierarchical SDMs has been proposed to address this issue by merging a global model calibrated with global, yet typically low-resolution, data and a local model calibrated with spatially restricted but more precise and reliable data[7] (Figure 1b).

Given the incredibly large number of species on Earth, it is essential to focus on sentinel species which can reflect the response of ecosystem health to environmental changes. Butterflies play a key ecological role, contributing to pollination, having a significant part in food webs[8], and hosting a great taxonomic diversity[9]. Butterflies are historically one of the most studied insects[10] and the recent involvement of the general public through citizen science initiatives[11] has generated a large number of observations worldwide.

ECEO

EPFL ENAC IIE
Devis Tuia, Prof.
Rue de l'Industrie 17
Case Postale 440
CH – 1951 Sion

Téléphone : +41 21 69 382 83 (secr).
E-mail : devis.tuia@epfl.ch

Project

This project aims to predict the distribution of butterflies globally and locally in Switzerland and Belize. The project will consist in two main parts:

• **GLOBAL**. Using butterfly observations worldwide from the Global Biodiversity Information Facility (GBIF, https://www.gbif.org/, with more than 145 millions of Lepidoptera observations, Figure 1a) and global climatic repositories (minimum, mean and maximum temperature, specific and relative humidity, evaporation, downwelling shortwave and longwave radiation, wind speed, precipitation, and cloud coverage) will be used to train and validate a deep learning SDM (Multilayer Perceptron) at the global scale.

•**LOCAL**. The global SDM outputs will be linked to estimate the local, actionable range estimations required by conservation actors through a hierarchical strategy (Figure 1). Switzerland and Belize will be used as emblematic study cases: Switzerland presents a temperate climate and a unique altitudinal gradient and is experiencing a biodiversity loss above the world average[12]. In contrast, Belize has a tropical climate and hosts irreplaceable ecosystems of global importance for species conservation and butterfly richness[13], which are particularly threatened by climate-change induced and anthropogenic pressures[14]. The Papillorama Foundation (https://www.papiliorama.ch/) represents a bridge between these two far-away countries, thanks to its conservation project in Belize, which established the Shipstern Nature Reserve.
Switzerland: High quality datasets of both species occurrence (Infospecies, https://www.infospecies.ch/de/) and climatic predictors (SWECO, https://zenodo.org/records/7994481) will be used to train and validate a national species distribution model of butterflies.
Belize: Species surveys carried out over the period 1990-1994 detected more than 120 species of butterflies in the Shipstern Nature Reserve. These data can be used to train the Belize national-scale SDM.

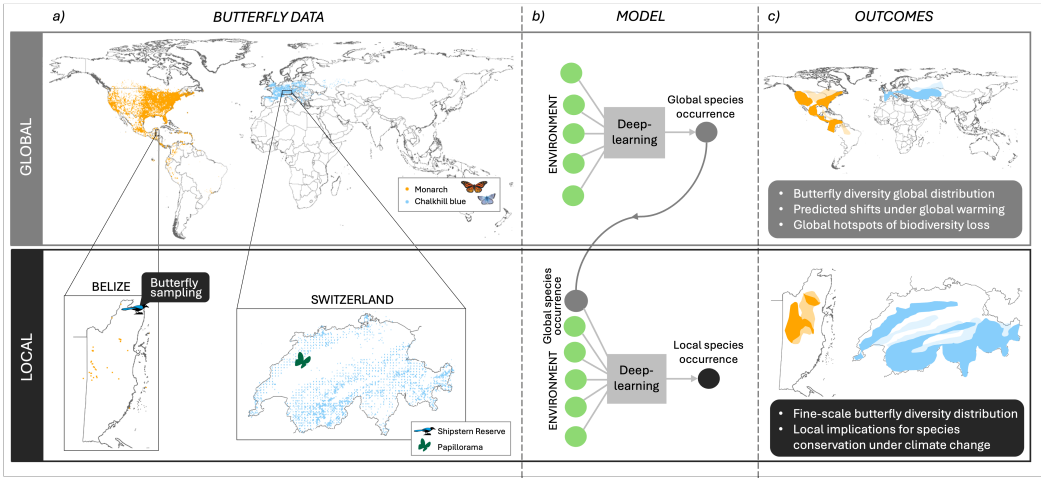

Figure 1. Global and local (Belize and Switzerland) examples of the Monarch butterfly (Danaus Plexippus, in orange) and Chalkhill blue butterfly (Lysandra coridon, in light blue). a) BUTTERFLY DATA: Global and local distribution of species occurrence data b) MODEL: Deep-learning modeling approaches, linking environmental covariates to species occurrence and embedding the global model output into the local-scale model; c) OUTCOMES: Expected species distribution maps.

ECEO

EPFL ENAC IIE
Devis Tuia, Prof.
Rue de l'Industrie 17
Case Postale 440
CH – 1951 Sion

Téléphone : +41 21 69 382 83 secr.
E-mail : devis.tuia@epfl.ch

Requirements

- Experience or strong interest in big data and modelling.
- Proficiency in Python and relevant libraries (e.g., pytorch, numpy, matplotlib).
- Strong willingness to learn and ability to work independently and interest in contributing to projects with ecological impact.

Literature

1. Watson, R. *et al.* Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. *IPBES Secr. Bonn Ger.* 22–47 (2019).
2. Guisan, A. *et al.* Predicting species distributions for conservation decisions. *Ecol. Lett.* **16**, 1424–1435 (2013).
3. Araújo, M. B. *et al.* Standards for distribution models in biodiversity assessments. *Sci. Adv.* **5**, eaat4858 (2019).
4. Brun, P. *et al.* Multispecies deep learning using citizen science data produces more informative plant community models. *Nat. Commun.* **15**, 4421 (2024).
5. Zbinden, R. *et al.* MaskSDM with Shapley values to improve flexibility, robustness, and explainability in species distribution modeling. Preprint at https://doi.org/10.48550/arXiv.2503.13057 (2025).
6. Cole, E. *et al.* Spatial implicit neural representations for global-scale species mapping. in *International conference on machine learning* 6320–6342 (PMLR, 2023).
7. Goicolea, T. *et al.* Spatially-nested hierarchical species distribution models to overcome niche truncation in national-scale studies. *Ecography* **2025**, e07328 (2025).
8. Ghazanfar, M., Malik, M. F., Hussain, M., Iqbal, R. & Younas, M. Butterflies and their contribution in ecosystem: A review. *J. Entomol. Zool. Stud.* **4**, 115–118 (2016).
9. Zhang, Z.-Q. *Animal Biodiversity: An Outline of Higher-Level Classification and Survey of Taxonomic Richness*. (Magnolia press, 2011).
10. New, T. R., Pyle, R. M., Thomas, J. A., Thomas, C. D. & Hammond, P. C. Butterfly Conservation Management. *Annu. Rev. Entomol.* **40**, 57–83 (1995).
11. Prudic, K. L. *et al.* eButterfly: leveraging massive online citizen science for butterfly conservation. *Insects* **8**, 53 (2017).
12. Explorateur des données de l'OCDE • Espèces menacées. https://data-explorer.oecd.org.
13. Shuey, J. A., Meerman, J., Labus, P., Kovarik, P. & Shuey, R. Butterfly community structure in belize: species richness patterns reveal the role of climate gradients in shaping butterfly community structure and strategies for long-term conservation. *J. Insect Conserv.* **28**, 233–249 (2024).
14. Baumbach, L., Hickler, T., Yousefpour, R. & Hanewinkel, M. Co-occurrence of climate-change induced and anthropogenic pressures in Central American key biodiversity areas. *Environ. Res. Lett.* **18**, 094035 (2023).

Contact
Chiara Vanalli chiara.vanalli@epfl.ch
Prof. Devis Tuia, devis.tuia@epfl.ch