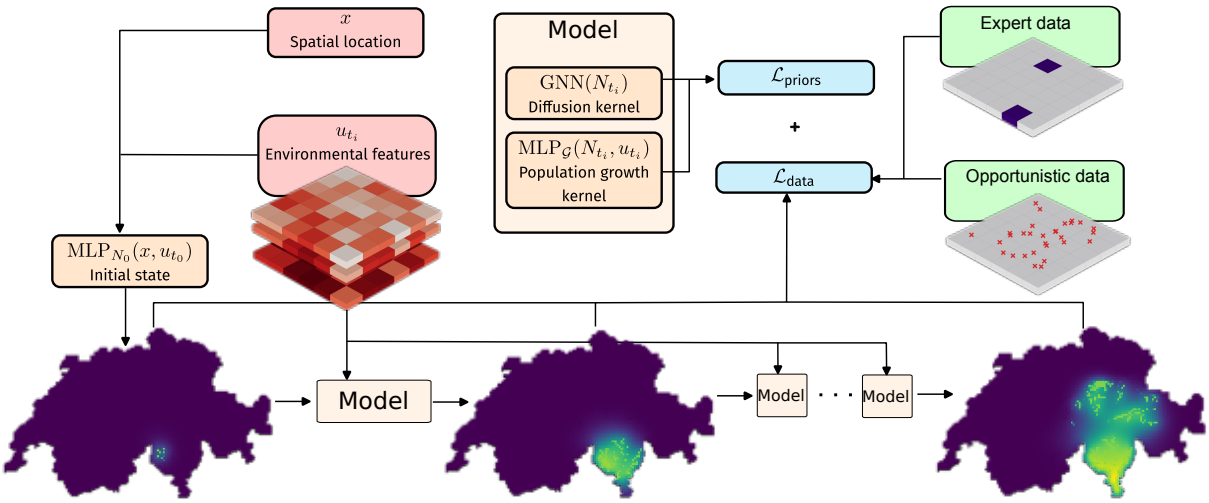
MSc thesis proposal

Forecasting invasive species range expansion using ecologically-informed neural networks

Supervisors: Michele Volpi[1], Konstantinos Tatsis[1], and Victor Boussange[2]

[1]SDSC, ETH Zurich
[2]Dynamic Macroecology Group, WSL

Spring Semester 2025



A deep-learning based reaction-diffusion model. The model is based on graph neural networks to model spatial dynamics coupled to simpler neural network to capture population growth. To make predictions, the model is used in an iterative roll-out fashion, predicting species range expansion in time. The initial state of the population is modeled by another neural network, estimating the state at time t_0. The model is constrained by ecological "priors", akin to physics-informed neural networks.

Project description

Background

Invasive Alien Species (IAS) are a primary driver of biodiversity loss, disrupt ecosystems, and impose substantial economic costs, exceeding $423 billion annually [1]. The number of IAS per country has increased by 70% since 1970 [1], and their continued spread threatens biodiversity and ecosystem services even in the absence of new introductions [2]. Despite the urgency of the issue, there remains a significant gap in predictive models capable of forecasting IAS range expansion [3]. This challenge comes from the complexity of invasion processes and limitations in existing datasets [4].

Modeling dynamics

Current state-of-the-art deep learning models can leverage exceptionally well large opportunistic datasets[1] to model species distributions [5]–[7]. However, these models assume that species are always in equilibrium with their environment, neglecting transient dynamics such as demography and range expansion, potentially leading to significant forecast errors [8].

Direct calibration of process-based models against time series data offers an alternative to purely data-driven methods. Process-based models explicitly represent such processes with two terms, one that relates to demographics and one to the diffusion in space in time [9]. They rely on strong ecological priors, making them theoretically more robust and data-efficient, while extrapolating under known conditions. However, they often struggle with inaccuracies and calibration challenges, reducing their ability to provide precise forecasts.

A novel approach?

"Physics-informed neural networks" (PINNs) [10] and related methods, commonly referred to as "scientific machine learning" methods [11], have recently emerged as a promising approach to bridge data-driven and process-based models. By constraining data-driven models with process-based relationshops, PINNs often demonstrate interesting features such as efficiency in data, generalization and credible outputs [10]–[12]. PINNs have been successful across various scientific fields, notably in climate and weather modeling [12], [13], but remain strongly underexplored for ecological applications [14], [15].

In parallel, Graph Neural Networks offer a framework to model irregularly sampled data in space and time, and can capture complex relationships at different spatial and temporal ranges [16]. Again, neural weather forecasting has shown great success when including this [13].

This project has potential to significantly improve conservation strategies by providing more reliable forecasts of IAS and transient responses of biodiversity to global change. To this end, it proposes the development of a novel, deep learning-based approach to model the evolution and expansion of invasive species in Switzerland. A first step is to explore the potential space of models and architecture (see initial Figure for an example), and train a constrained making use of synthetic data describing species range expansion as generated by a differential equation. In a second step, we will focus on assimilating opportunistic, high-volume, presence only data.

[1] We refer to collections of data that are gathered without a predefined plans or structured approaches to answer specific questions. These datasets are often derived as a byproduct of other activities or processes.

Research Gaps

- Can we obtain data-driven dynamical predictions of species range expansion by explicitly taking into account population and spatial dynamics?
- How can we assimilate presence-only observations with sampling biases from opportunistic datasets?

Project Goals

- Identify and implement a relevant architecture for the core dynamical model.
- Evaluate and validate the performance of the framework in an ideal data regime.
- Develop an appropriate data assimilation scheme that can leverage realistic datasets with sampling biases.
- Create and setup a code repository allowing experiments to be reproduced, models extended, and source maintained by other collaborators.

Your qualifications

- Keen interest in dynamical systems.
- Experience with graph neural networks.
- Strong programming skills in Python, with experience in JAX and/or Pytorch (PyG, Lightning).
- Experience with version control systems (Git)

What we offer

- An innovative project with large societal and scientific impact.
- The opportunity to publish you work in a top-tier ML conference to boost up your CV.
- Close supervision and mentoring by experienced researchers.
- Collaboration with experts from world-renowned institutions and access to cutting-edge research facilities.
- Periodically available desk space at SDSC offices in Andreasturm.

Contact

We are delighted to hear from you! Please contact Michele via e-mail to michele.volpi@sdsc.ethz.ch. We will be happy to answer any question and set up a call to discuss the project.

Bibliography

[1] H. E. Roy, A. Pauchard, P. Stoett, *et al.*, "IPBES Invasive Alien Species Assessment: Summary for Policymakers," Zenodo, Tech. Rep., Sep. 2023.

[2] H. Seebens, T. M. Blackburn, E. E. Dyer, *et al.*, "No saturation in the accumulation of alien species worldwide," *Nature Communications*, vol. 8, no. 1, p. 14 435, Feb. 2017.

[3] B. Lenzner, D. Leclère, O. Franklin, H. Seebens, N. Roura-Pascual, M. Obersteiner, S. Dullinger, and F. Essl, "A Framework for Global Twenty-First Century Scenarios and Models of Biological Invasions," *BioScience*, vol. 69, no. 9, pp. 697–710, Sep. 2019.

[4] H. E. Roy, A. Pauchard, P. Stoett, and T. Renard Truong, "IPBES Invasive Alien Species Assessment: Full report," Zenodo, Tech. Rep., Feb. 2024.

[5] P. Brun, D. N. Karger, D. Zurell, P. Descombes, L. C. de Witte, R. de Lutio, J. D. Wegner, and N. E. Zimmermann, "Multispecies deep learning using citizen science data produces more informative plant community models," *Nature Communications*, vol. 15, no. 1, p. 4421, May 2024.

[6] E. Cole, G. Van Horn, C. Lange, A. Shepard, P. Leary, P. Perona, S. Loarie, and O. Mac Aodha, *Spatial Implicit Neural Representations for Global-Scale Species Mapping*, Jun. 2023.

[7] L. E. Gillespie, M. Ruffley, and M. Exposito-Alonso, "Deep learning models map rapid plant species changes from citizen science and remote sensing data," *Proceedings of the National Academy of Sciences*, vol. 121, no. 37, e2318296121, Sep. 2024.

[8] M. C. Urban, "Accelerating extinction risk from climate change," *Science*, vol. 348, no. 6234, pp. 571–573, May 2015.

[9] M. Bonneau, F. A. Johnson, and C. M. Romagosa, "Spatially explicit control of invasive species using a reaction–diffusion model," *Ecological Modelling*, vol. 337, pp. 15–24, Oct. 2016.

[10] M. Raissi, P. Perdikaris, and G. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *Journal of Computational Physics*, vol. 378, pp. 686–707, Feb. 2019.

[11] C. Rackauckas, Y. Ma, J. Martensen, C. Warner, K. Zubov, R. Supekar, D. Skinner, and A. Ramadhan, "Universal Differential Equations for Scientific Machine Learning," *Manuscript*, vol. XXX, no. Xx, pp. 1–8, Jan. 2020.

[12] K. Kashinath, M. Mustafa, A. Albert, *et al.*, "Physics-informed machine learning: Case studies for weather and climate modelling," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 379, no. 2194, 2021.

[13] R. Lam, A. Sanchez-Gonzalez, M. Willson, *et al.*, *GraphCast: Learning skillful medium-range global weather forecasting*, Dec. 2022.

[14] J. H. Lagergren, J. T. Nardini, R. E. Baker, M. J. Simpson, and K. B. Flores, "Biologically-informed neural networks guide mechanistic modeling from sparse experimental data," *PLOS Computational Biology*, vol. 16, no. 12, e1008462, Dec. 2020.

[15] A. Yazdani, L. Lu, M. Raissi, and G. E. Karniadakis, "Systems biology informed deep learning for inferring parameters and hidden dynamics," *PLOS Computational Biology*, vol. 16, no. 11, V. Hatzimanikatis, Ed., e1007575, Nov. 2020.

[16] P. Veličković, *Everything is Connected: Graph Neural Networks*, Jan. 2023.