



Forward and inverse modelling of eco-evolutionary dynamics



in biological and economic systems

Victor BOUSSANGE

Victor Boussange

Forward and inverse modelling of eco-evolutionary dynamics in biological and economic systems

September 13, 2022

Cover picture: Top: forest in Sorapiss, Dolomites, Italy. Bottom: New York City, USA. @ Luca
Bravo / PhotoSpirit

The document format is based on the *Clean Thesis* style developed by Ricardo Langner.

Contents

1	Introduction	1
2	Discussion	15
2.1	Contributions	15
2.1.1	Advances in the fundamental understanding of biological and economic systems	15
2.1.2	Methodological advances in forward and inverse modelling .	19

Introduction

„ Nature Loves to Hide.

— Heraclitus

Biological and economic systems as complex adaptive systems

What are the similarities between biological and economic systems? Both are complex adaptive systems (CAS) [1], which are composed of heterogeneous entities structured at different levels of organizations, that interact in nonlinear ways and experience evolutionary processes. Interaction and evolutionary processes take many different forms and operate at different organizational level [Levin1998] (see Fig. 1.1). Interestingly, the variety of processes involved and their couplings do not necessarily lead to unpredictable, chaotic, or erratic structures and dynamics [Olff2009], but rather induce organised structural properties and behavior [mitchell2009complexity]. In biological systems, those include patterns of species richness, where for instance montane regions are often associated with a disproportionately high number of species [2]. In economic systems, those include the distribution of international income, where some countries have systematically developed much more rapidly than others [acemoglu2001colonial]. A common direction on the research agenda in Biology and Economics is to comprehend the set of interaction and evolutionary processes that determine these emergent properties [3], and how do they do so. In biological systems, the nature of the processes of interaction and evolution is identified, and the current challenge is to comprehend the mechanisms resulting from their couplings. In economic systems, we still do not exactly understand the nature of those processes, and how are they involved.

Ecological and evolutionary processes drive the dynamics of biological systems

In biological systems, interaction processes are generally designated as ecological processes, and involve fluxes of energy and matter across space and time, encompassing the processes of interaction between organisms (biotic interactions) and between organisms and their environment (abiotic interactions), and dispersal processes (movement of individual across space) [4]. Evolutionary processes designate those

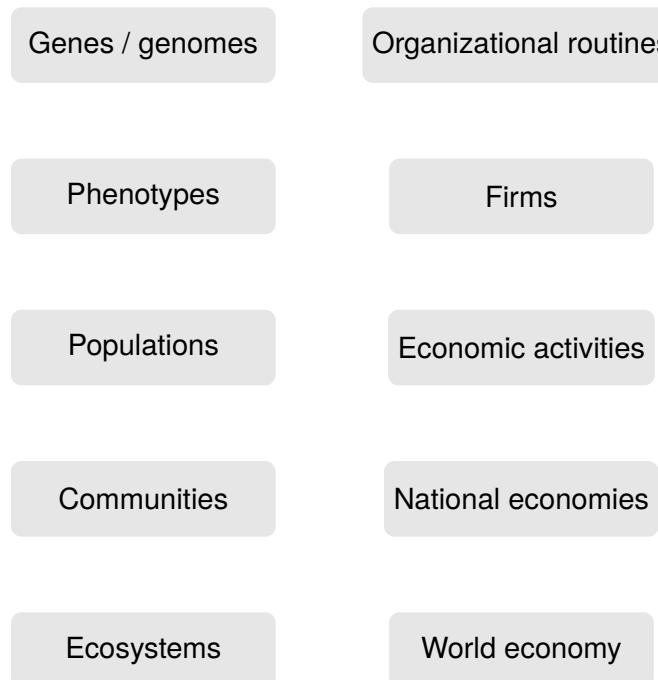


Fig. 1.1: Schematic diagram of proposed organisational levels in biological and economic systems. A is inspired from [Hendry+2016]

processes responsible for the change of heritable characteristics (DNA, genes, phenotypes) over successive generations [Hall2013]. The coupling between ecological and evolutionary processes is acknowledged since the very birth of the theory of Evolution, when Darwin realised a link between the different ecological opportunities across the Galápagos islands and the different beak shapes in the finches he found in each island [darwin2004origin], during his voyage on the Beagle. He reasoned that the variations in ecological opportunities lead to a differential in survival for certain phenotypes, which over time resulted in the evolution of different beak shapes. Since then, we know that ecological processes directly affect evolutionary response [Ezard2009]. In the recent years, the idea that not only ecological processes can affect evolutionary response, but also that evolutionary processes could affect ecological processes, has developed [XXX]. Empirical studies have demonstrated that evolution can be rapid and occur on similar time scales as ecology [Hairston2005, 5] and have quantifiable effects on ecological dynamics [Ezard2009], leading to feedbacks between ecological and evolutionary processes, so-called eco-evolutionary feedbacks [Schoener2011, 5]. Eco-evolutionary feedbacks involve situations where an ecological property influences evolutionary change, which then feeds back to an ecological property, or vice versa [Govaert2019a]. Examples are feedbacks between population density (ecological property) and trait evolution (evolutionary

change), which can lead to evolutionary branching through the effect of competition [Doebeli1999]. Eco-evolutionary feedbacks are also involved in adaptation mechanisms [Doebeli1999], where species disperse and phenotypic variations allow to adapt to local environments [XXX]. Those feedbacks may greatly influence the mechanisms driving the dynamics of ecosystems [6], but our understanding of their nature and effect is limited [Lion2022]. In particular, eco-evolutionary feedbacks are expected to play a critical role in the evolution of the biosphere in the coming decades [7], as ecosystems are being rapidly affected by anthropogenic pressure and with climate change [8, 9]. In order to mitigate the consequences of human development, it is of utmost urgency to better understand eco-evolutionary feedbacks [7], and develop mechanistic models embedding this knowledge [6]. This will in turn provide more reliable forecasts of ecosystem states [Clark2001], to help designing adequate management of ecosystem services [6].

Drivers of economic change

The processes that determine economic change is controversial in economics [Nelson2014]. To explain economic development, mainstream economic theory [10.1093/cje/bet027] assumes that economic systems are in equilibrium, in the sense that the demand and supply of goods and services are balanced on all relevant markets. Firms are rational in maximizing profits by adapting to demand and supply, and the observed economic change is driven by exogenous forces, such as technological change [Romer1986]. Evolutionary economics, promoted by the seminal work of Ref. [Nelson2014], criticizes this view and seeks to explain economic change by focusing on endogenous forces. Evolutionary economics suggests that interactions between firms and economic activities, and evolutionary processes acting upon them, are major processes contributing to economic change [Hodgson2019]. For instance, firms or economic activities may interact positively or negatively [Wernerfelt1989, Pistorius2007Ozman2009, Saavedra2009a, Cohendet2018, Menon2015], spread across space [RogersEverettM2003DoI, Zahra2000], and adapt [Cordes2006] or transform into new economic institutions [Freeman2002, Hodgson2004, Aldrich2008], affecting economic development at the regional and national scale. Because these processes are analogous to eco-evolutionary processes driving the dynamics of biological systems, a number of modelling approaches have borrowed concepts and methods from biology in the last decades, aiming at better understanding the processes underlying emergent properties in economic systems [Tacchella2018, Saavedra2009a, Scholl2020, Zhang2018, Modis1997, Saavedra2014, Farmer1999, Michalakelis2011, Marasco2016, Gatabazi2019, Cauwels56, Applegate2021, Suweis2015]. For instance, [Saavedra2009a] has

successfully used a model of mutualistic interaction to explain structural patterns in industrial cooperation. Also, [Scholl2020] uses the concepts of foodwebs and density dependence to explain market malfunctions and excess volatility in financial markets. However, those studies did not seek to understand how these processes may affect economic development. Recent modelling approaches developed in evolutionary biology may help to disentangle whether eco-evolutionary processes could explain differences in economic development across countries.

Forward modelling of eco-evolutionary processes

The complex interplay between ecological and evolutionary processes, acting at different scales of time and space and organization, can hardly be studied with experimental approaches [Hagen2022, 10]. As such, a deductive approach, relying on forward modelling, has traditionally been put forward to advance our understanding of the mechanisms underlying [Brummitt2020]. Along this approach, hypotheses about causal processes are embedded in a model, which forward integration generates emergent properties. Such emergent behavior may be seen as predictions from the consideration of the causal processes [May2004]. The role of the modeller is to point at the mechanisms by which the properties emerge, disentangling the underlying interplay between the processes. In the early 1930s to 1940s, by formulating tractable mathematical models implementing the processes of reproduction, dispersal and mutations, the work of Fisher, Wright and Haldane has greatly contributed to the modern synthesis of evolutionary biology [huxley1942evolution], generally accepted as the basis of our current understanding of evolutionary dynamics. The mechanistic models commonly take the form of differential equations (DE), and express how the processes under investigation affect the rate of change of the population characteristics, such as the proportion of a given allele. However, the requirement of tractable mathematical models (DEs that yield analytical solutions) has involved strong assumptions on the processes investigated, that are poorly representative of the complexity of eco-evolutionary feedbacks in nature [Govaert2019a]. In particular, ecological scenarios have been strongly simplified, and did not take into account how evolution could affect population dynamics [Lion2022]. As such, traditional mathematical models have omitted eco-evolutionary feedbacks and density dependence.

With the increase in computational capacity, novel modelling approaches relying on individual based models (IBMs) have appeared [deangelis2005individual]. IBMs allow to capture processes acting at the individual level, requiring less simplifying assumptions than traditional mathematical models [deangelis2005individual]. Capturing more realistic scenarios by allowing the forward integration of com-

plex hypothesis, the lack of analytical tractability of IBMs may nonetheless occult the mechanisms underlying emergent properties [May2004, 11]. The recent development of mathematical techniques, such moment closure approximations [law1999moment, Gandhi2000, 12, 11], adaptive dynamics theory [Metz1995], and probability theory [13], are generating novel pathways by filling the gap between IBMs and mathematical models. Analogous to renormalisation group analysis developed in quantum and statistical physics [Sayama], they form a toolbox to rigorously derive how emergent properties are influenced by processes operating at different organizational levels. As such, they allow an analytical underpinning to IBM simulations, and can generate a general understanding of the key mechanisms at stake [11].

The combination of numerical simulations and, e.g., adaptive dynamics theory, has successfully shed new lights on the emergence of evolutionary branching under frequency-dependent selection [14, 15]. An other example is the work of [Debarre2013a, 16, 17], that has provided new insights on the effect of habitat heterogeneity on population dynamics. However, our current understanding of eco-evolutionary feedbacks omits potentially significant factors, such as the structuration of populations over complex spatial structures [XX] and highly dimensional phenotypic space [XXX].

The consideration of such details is important to advance our understanding, but raises challenging methodological issues. In particular, complex models may hinder the fundamental mechanisms underlying the emergence of a pattern. Also, the consideration of multiple traits leads to an increase in the dimensionality of the associated DE problem, which in turn leads to an exponential increase in computational cost [XXX]. In order to better understand eco-evolutionary feedbacks, we need to investigate more realistic scenarios, which will, in turn, require the development of novel numerical methods that can cope with the extra computational cost.

Inverse modelling

While in forward modelling, observation data is uniquely considered at the end of the modelling cycle - when comparing predicted and empirical patterns to refine the baseline hypotheses -, inverse modelling integrates data at the start the modelling cycle [XXX](??). Inverse modelling can take the form of parameter estimation [18] or model selection [19], both involving the use of inference methods to estimate, respectively, the most probable model parameter value, or the most probable model among candidates, given the data. Provided that they are inferred together with uncertainties, parameters can be interpreted to better understand the strengths and

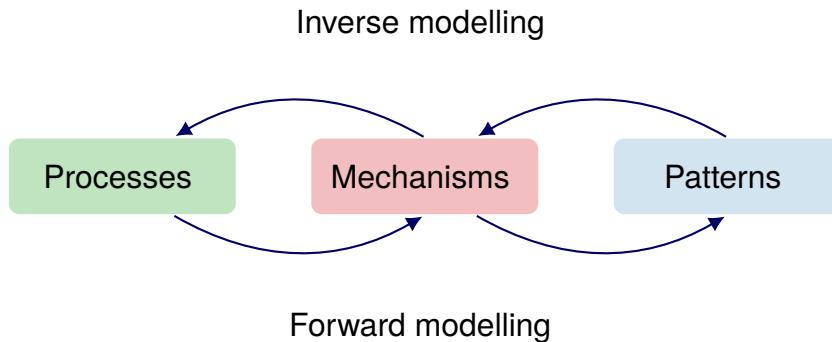


Fig. 1.2: Forward and inverse modelling approaches for the understanding of CAS. A forward modelling approach starts with assuming a set of processes, and seeks to understand how their interplay transforms in mechanisms that are associated with macroscopic patterns. An inverse modelling approach starts from empirical patterns (or observations) to deduce the underlying mechanisms and associated processes.

effects of the associated processes [10]. For instance, [20, 21] infer and analyse the parameters of population dynamic models to understand processes involved in ecosystem functions. In model selection, the candidate models embed competing hypotheses about processes and mechanisms, and the relative support of each model given the data is used to discriminate [19]. For instance, [Skeels2022] shows that temperature-dependent evolutionary speed is the most likely mechanism to explain variations in biodiversity patterns, using inference methods to discriminate between alternative dynamic models embedding different evolutionary hypotheses. The inference methods used to calculate the parameter and model probabilities proceed by defining a distance between the model simulation and the observation data [XXX], which relates to the probability of the parameters given the model and the data [18]. In parameter estimation, the most likely parameters are associated with the minimum distance to the data [XXX], and in model selection, the most likely models are associated with the minimum of the minimum model distance [XXX]. Inference methods involve numerical methods to recover the model distance to the data [XXX], which success is dependent on the model formulation and its dynamical properties [XXX]. Thus far, the use of inverse modelling for understanding eco-evolutionary dynamics has been limited because of a number of issues, including the high computational cost of the forward integration of eco-evolutionary models [22], the large number of parameters involved [23], and their strong nonlinearities [24, 25, 26]. Advances in the field of artificial intelligence could circumvent these issues, allowing to confront eco-evolutionary models with data [XXX].

Machine learning to leverage forward and inverse modelling

In the recent years, the field of artificial intelligence (AI) has made enormous progresses in computer vision [XXX] and natural language processing [XXX]. At the backbone of this success are key computational techniques that could leverage the forward and inverse modelling of CAS. Computer vision and natural language processing rely on deep learning methods, that allow neural networks to learn abstract representation of mechanisms from large datasets [27]. These abstractions are hardly interpretable by humans [XXX], and their prediction ability is limited by the information contained in the training datasets. As such, neural networks cannot be used *per se* to gain scientific insights and extrapolate beyond observed trends [28, 6]. Nevertheless, their traditional applications and associated methods have been successfully derived in other scientific fields for this purpose [Rolnick2023, 29, 30, 31].

Neural networks have been used in forward modelling, to reduce the cost of the forward integration of climate models, by learning more efficient representations of physical mechanisms [XXX]. They have also been used to approximate the solution of partial differential equations (PDEs) [32], with the major advantage of approximating high dimensional problems at a lower computational cost than traditional methods. Underlying the training of neural network is the technique of backpropagation [XXX]. This technique can be generalised to any scientific model against data [33], with the potential to leverage inverse modelling techniques. As such, the derivation of AI techniques to investigate causal processes in CAS offers unique opportunities [Frank2022].

Programming languages

Combining ML techniques with scientific models requires computational environments that allow to easily develop scientific models, ensure simulation performance, and provide composability between ML and other scientific libraries [33]. Unfortunately, performance and composability are features that are poorly represented in mainstream programming languages used by the scientific community, such as Python, Matlab or R. Those languages are naturally attractive because they are dynamically typed [XX], allowing convenient development iterations. Nonetheless, prototypes written in Python, Matlab or R need to be rewritten in low level, compiled languages such as C, C++ or Fortran for speed and predictable mapping to hardware [Perkel2019, Bezanson2017]. This conversion requires significant involvement, leading to a problem commonly designated as the "two language problem" [Bezanson2017]. In order to circumvent issues of performance, most libraries in Python, Matlab or R rely on bindings with low level languages. For instance, the

most used deep learning libraries in Python, TensorFlow and PyTorch, are internally written in C. However, bindings with low level languages come with major negative externalities, such as restricting the understandability of their internals to computer scientists – prohibiting potential development contributions from the scientific community –, and preventing the composability of libraries [XXX].

Julia is a programming language that was launched in 2012 to address the issue of the two-language problem [Bezanson2017, Bezanson2018]. Julia was built over a type-specializing, just-in-time compiler, which makes it easy to generate highly performant programs, while preserving the essential features of Python, Matlab or R, such as dynamic typing and automatic memory management. Importantly, it relies on multiple dispatch, which allows to generate highly generic code with good performance. This permits to write libraries in pure Julia, guaranteeing productivity and composability. As such, the internals of any Julia library can be understood by non computer-scientists, who can further use his expertise to participate to its development. Many Julia libraries benefit from a high number of contributions of independent users (see, e.g., github.com/DifferentialEquations.jl). Multiple dispatch also allows to automatically generate the gradient of any Julia program without any modification [ForwardDiff.jl, Zygote.jl]. This means that any scientific library in Julia, such as differential equation solvers, can be combined with deep learning tools, with unique opportunities for forward and inverse modelling problems [Frank2022]. As such, Julia allows to prototype a program which is readily generic and can directly be shared to the research community. Further, by granting the composability of libraries, it allows to blend ML techniques with scientific models. This makes Julia a promising computational environment to accelerate research in CAS.

Thesis outline

In summary, while it is increasingly acknowledged that feedbacks between ecological and evolutionary processes play an important role in the dynamics of biological systems [5, 6], our understanding of the mechanisms in which they are involved has been limited to simplified scenarios. Further, while analogous processes have been suggested to influence the dynamics of economic systems [Hodgson2019], a quantification of their effect is missing. Under increasing anthropogenic pressure, these research directions become essential [6], but raise challenging methodological issues. Here, I present novel forward and inverse modelling approaches to advance our understanding of eco-evolutionary dynamics in biological and economic systems, and utilise them to shed light on the underlying processes and resulting mechanisms.

In ??, I investigate how eco-evolutionary processes, in combination with complex habitat spatial structures, influence the trait distribution of biological populations. I

proceed using a forward modelling approach, building a stochastic eco-evolutionary IBM where individuals are structured over a spatial graph, and experience the fundamental processes of reproduction, competition, mutation and migration. I seek to understand how those microscopic forces result in trait differentiation at the population level. I derive DE approximations of the IBM that, together with extensive numerical simulations, provide analytical insights into how the graph properties affect the population size and trait differentiation. In particular, I show that three main graph properties, measuring landscape connectivity, heterogeneity in connectivity, and habitat spatial auto-correlation, shape the trait differentiation of the biological population. These results establish mechanistic links between landscape features and the eco-evolutionary dynamics of biological populations.

In ??, I develop an inverse modelling method to estimate the parameters of highly nonlinear population dynamic models. The method is based on a machine learning framework and involves AI techniques together with a novel learning strategy. This learning strategy consists in training the model against mini-batches of data with short time horizon, which I analytically show to regularize the learning problem. I implement the ML framework in the Julia library **MiniBatchInference.jl**, and demonstrate through numerical experiments that it can efficiently estimate model parameters and provide statistical evidences for causal processes from noisy, incomplete and independent time series. Altogether, the proposed ML framework is a workhorse for inverse modelling and can elucidate mechanistic pathways in biological and economic systems.

In ??, I quantify the effect of eco-evolutionary processes on the dynamics of economic systems. I employ the ML framework developed in ?? to investigate how alternative eco-evolutionary population dynamic models can explain the dynamics of economic activities in the richest 100 countries, relying on 59 year of economic data. The models embed the processes of ecological interactions between economic activities, spatial transfers, and economic activity transformations, which support is compared to a simple logistic growth model, taken as a null model. I find strong statistical evidence for positive interactions between national economic activities, and spatial transfers across countries. To my knowledge, this is the first study that provides quantitative evidences that similar processes may influence the dynamics of biological and economic systems.

In ??, I extend two recent methods to solve high dimensional PDEs, in order to handle non-local nonlinear terms. The first method relies on Picard iterations, while the second is based on machine learning and involves neural networks to approximate PDE solutions. The numerical difficulties arising due to the non-local term are avoided by using a plain vanilla Monte Carlo integration. I implement the methods in the Julia library **HighDimPDE.jl**, and evaluate their performance on

high dimensional PDE models arising in physics and biology, including population dynamic eco-evolutionary models. For all models, the methods yield good results with short run times, offering the possibility to include more realism in future eco-evolutionary models.

References

- [1] S. A. Levin. “Complex adaptive systems: Exploring the known, the unknown and the unknowable”. In: *Bulletin of the American Mathematical Society* 40.01 (2002), pp. 3–20. DOI: 10.1090/S0273-0979-02-00965-5.
- [2] C. Rahbek et al. “Humboldt’s enigma: What causes global patterns of mountain biodiversity?” In: *Science* 365.6458 (2019), pp. 1108–1113. DOI: 10.1126/science.aax0149.
- [3] J. M. Nordbotten et al. “Ecological and evolutionary dynamics of interconnectedness and modularity”. In: *Proc. Natl. Acad. Sci. USA* 115.4 (2018), pp. 750–755. DOI: 10.1073/pnas.1716078115.
- [4] M. Vellend. “Conceptual Synthesis in Community Ecology”. In: *The Quarterly Review of Biology* 85.2 (2010), pp. 183–206. DOI: 10.1086/652373.
- [5] F. Pelletier, D. Garant, and A. Hendry. “Eco-evolutionary dynamics”. eng. In: *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 364.1523 (2009), pp. 1483–1489. DOI: 10.1098/rstb.2009.0027.
- [6] M. C. Urban et al. “Improving the forecast for biodiversity under climate change”. In: *Science* 353.6304 (2016). DOI: 10.1126/science.aad8466.
- [7] J. Norberg et al. “Eco-evolutionary responses of biodiversity to climate change”. In: *Nature Climate Change* 2.10 (2012), pp. 747–751. DOI: 10.1038/nclimate1588.
- [8] E. C. Ellis. “Anthropogenic transformation of the terrestrial biosphere”. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 369.1938 (2011), pp. 1010–1035. DOI: 10.1098/rsta.2010.0331.
- [9] G. Midgley and L. Hannah. “Extinction Risk from Climate Change”. In: *Biodiversity and Climate Change*. Yale University Press, 2019, pp. 294–296. DOI: 10.2307/j.ctv8jnzw1.37.
- [10] M. Pontarp, Å. Bränström, and O. L. Petchey. “Inferring community assembly processes from macroscopic patterns using dynamic ecoevolutionary models and Approximate Bayesian Computation (ABC)”. In: *Methods in Ecology and Evolution* 10.4 (2019). Ed. by T. Poisot, pp. 450–460. DOI: 10.1111/2041-210X.13129.
- [11] S. Lion. “Moment equations in spatial evolutionary ecology”. In: *Journal of Theoretical Biology* 405 (2016), pp. 46–57. DOI: 10.1016/j.jtbi.2015.10.014.

- [12] J. M. Nordbotten et al. “The dynamics of trait variance in multi-species communities”. In: *R. Soc. Open Sci.* 7.8 (2020), Article No. 200321, 20 pp. DOI: 10.1098/rsos.200321.
- [13] N. Champagnat, R. Ferrière, and S. Méléard. “Unifying evolutionary dynamics: From individual stochastic processes to macroscopic models”. In: *Theoretical Population Biology* 69.3 (2006), pp. 297–321. DOI: 10.1016/j.tpb.2005.10.004.
- [14] U. Dieckmann and M. Doebeli. “On the origin of species by sympatric speciation”. In: *Nature* 400.6742 (1999), pp. 354–357. DOI: 10.1038/22521.
- [15] M. Doebeli and U. Dieckmann. “Speciation along environmental gradients”. In: *Nature* 421.6920 (2003), pp. 259–264. DOI: 10.1038/nature01274.
- [16] G. Meszéna, I. Czibula, and S. Geritz. “Adaptive Dynamics in a 2-Patch Environment: A Toy Model for Allopatric and Parapatric Speciation”. In: *Journal of Biological Systems* 05.02 (1997), pp. 265–284. DOI: 10.1142/S0218339097000175.
- [17] S. Mirrahimi and S. Gandon. “Evolution of specialization in heterogeneous environments: equilibrium between selection, mutation and migration”. In: *Genetics* 214.2 (2020), pp. 479–491. DOI: 10.1534/genetics.119.302868.
- [18] M. Schartau et al. “Reviews and syntheses: parameter identification in marine planktonic ecosystem modelling”. In: *Biogeosciences* 14.6 (2017), pp. 1647–1701. DOI: 10.5194/bg-14-1647-2017.
- [19] J. B. Johnson and K. S. Omland. “Model selection in ecology and evolution”. In: *Trends in Ecology & Evolution* 19.2 (2004), pp. 101–108. DOI: 10.1016/j.tree.2003.10.013.
- [20] S. I. Higgins, S. Scheiter, and M. Sankaran. “The stability of African savannas: insights from the indirect estimation of the parameters of a dynamic model”. In: *Ecology* 91.6 (2010), pp. 1682–1692. DOI: 10.1890/08-1368.1.
- [21] A. Curtsdotter et al. “Ecosystem function in predator-prey food webs—confronting dynamic models with empirical data”. In: *Journal of Animal Ecology* 88.2 (2019). Ed. by D. Stouffer, pp. 196–210. DOI: 10.1111/1365-2656.12892.
- [22] R. A. Fisher et al. “Vegetation demographics in Earth System Models: A review of progress and priorities”. In: *Global Change Biology* 24.1 (2018), pp. 35–54. DOI: 10.1111/gcb.13910.
- [23] I. L. Boyd. “The Art of Ecological Modeling”. In: *Science* 337.6092 (2012), pp. 306–307. DOI: 10.1126/science.1225049.

- [24] A. Hastings et al. “Chaos in Ecology: Is Mother Nature a Strange Attractor?” In: *Annual Review of Ecology and Systematics* 24.1 (1993), pp. 1–33. DOI: 10.1146/annurev.es.24.110193.000245.
- [25] J. Huisman and F. J. Weissing. “Biodiversity of plankton by species oscillations and chaos”. In: *Nature* 402.6760 (1999), pp. 407–410. DOI: 10.1038/46540.
- [26] E. Benincà et al. “Chaos in a long-term experiment with a plankton community”. In: *Nature* 451.7180 (2008), pp. 822–825. DOI: 10.1038/nature06512.
- [27] Y. LeCun, Y. Bengio, and G. Hinton. “Deep learning”. In: *Nature* 521.7553 (2015), pp. 436–444. DOI: 10.1038/nature14539.
- [28] A. D. Barnosky et al. “Approaching a state shift in Earth’s biosphere”. In: *Nature* 486.7401 (2012), pp. 52–58. DOI: 10.1038/nature11018. arXiv: 9605103 [cs].
- [29] K. Kashinath et al. “Physics-informed machine learning: Case studies for weather and climate modelling”. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 379.2194 (2021). DOI: 10.1098/rsta.2020.0093.
- [30] T. Schneider et al. “Earth System Modeling 2.0: A Blueprint for Models That Learn From Observations and Targeted HighResolution Simulations”. In: *Geophysical Research Letters* 44.24 (2017), pp. 12,396–12,417. DOI: 10.1002/2017GL076101. arXiv: 1709.00037.
- [31] A. Yazdani et al. “Systems biology informed deep learning for inferring parameters and hidden dynamics”. In: *PLOS Computational Biology* 16.11 (2020). Ed. by V. Hatzimanikatis, e1007575. DOI: 10.1371/journal.pcbi.1007575.
- [32] J. Sirignano and K. Spiliopoulos. “DGM: A deep learning algorithm for solving partial differential equations”. In: *J. Comput. Phys.* 375 (2018), pp. 1339–1364.
- [33] C. Rackauckas et al. “Universal Differential Equations for Scientific Machine Learning”. In: *arXiv:2001.04385v3* (2020), 18 pages.

Discussion

„ Les données pertinentes détiennent les réponses.

— French anagram

2.1 Contributions

Understanding biological and economic systems involves the understanding of the necessary and sufficient elemental processes, and their couplings, that result in mechanisms constraining their dynamics [1]. Along this line, my work aimed at advancing our quantitative understanding of how ecological and evolutionary processes, and their interplay, shape the dynamics of biological and economic systems. In particular, this thesis contributed to

- (i) a fundamental understanding of the role of eco-evolutionary processes in shaping the dynamics of biological populations structured in complex landscapes [chap1],
- (ii) the quantification of the effect of mechanisms associated with eco-evolutionary processes in economic systems ??,
- (iii) methodological advances in the forward and inverse modelling of eco-evolutionary dynamics ??????.

In the following, I discuss the chapters of my thesis collectively within the broader context of our current understanding of the dynamics of biological and economic systems, and the accompanying modelling paradigm.

2.1.1 Advances in the fundamental understanding of biological and economic systems

Linking eco-evolutionary processes to patterns of differentiation

Spatial patterns of biodiversity result from microscopic processes acting upon individual organisms [2], which include reproduction, competition, mutation and dispersal. Mutations result in genetic drift, which promotes stochastic variations in the allelic proportions and phenotypes of biological populations [XXX]. In spatially structured

populations, this results in turn to "neutral differentiation", where spatially isolated populations are characterised by differentiated alleles and traits [XXX]. Dispersal tends to reduce neutral differentiation, and this effect is modulated by landscape connectivity [3, 4, 5] through the mechanism of "isolation by limited dispersal" [6]. By increasing the dispersal ability of organisms, landscape connectivity decreases neutral differentiation. When landscapes present heterogeneous habitats, natural selection can supplement the effect of genetic drift and increase the sole effect of stochasticity on differentiation. Under this scenario, local environment conditions select individuals with traits that provide them higher fitness [XXX]. At the population level, this results in populations adapting to their local environment, a mechanism coined "local adaptation" [7] and resulting in patterns of "adaptive differentiation". Adaptive differentiation is hindered by dispersal, which prevents local adaptation by bringing maladapted individuals, that destabilise the evolution of traits towards the optimal. While adaptive differentiation concerns traits under selection, it indirectly affects the differentiation of neutral traits, that are co-evolving with traits under selection through linkages [XXX]. This results in turn to the mechanism of "isolation by adaptation", where habitat heterogeneity, rather than landscape connectivity, increases neutral differentiation [nosil2008]. Simple mechanisms resulting in neutral and adaptive differentiation are identified, but how they are modulated by eco-evolutionary feedbacks and landscape complexity is unclear.

In ??, I demonstrate a novel mechanism, involving the ecological process of competition, that considerably affects neutral differentiation. Through the creation of unbalanced migration fluxes which increases the intensity of competition in highly connected populations, heterogeneity in connectivity reduces gene flow and reinforces neutral differentiation. I also investigate the mechanism of local adaptation and how it results in adaptive differentiation in complex landscapes, where habitats are arranged in a realistic fashion. I show that the complexity of habitat spatial distribution can be reduced to a measure of habitat spatial auto-correlation, coined the "habitat assortativity". Landscapes characterised by a high habitat assortativity systematically support populations that are locally better adapted than in landscape with low assortativity, resulting in higher adaptive differentiation. Specifically, I provide an analytical condition for local adaptation that sheds light on how it relates to dispersal intensity, selection strength, habitat heterogeneity, and habitat assortativity.

Because habitat assortativity affects local adaptation, it must also affect neutral differentiation through the mechanism of isolation by adaptation. Closing the loop, I demonstrate that habitat assortativity affects population differentiation through two antagonistic effects. By favoring local adaptation, it promotes isolation by adaptation, therefore increasing neutral differentiation. In parallel, it also favors gene flow within clusters of similar environmental conditions, decreasing isolation by limited

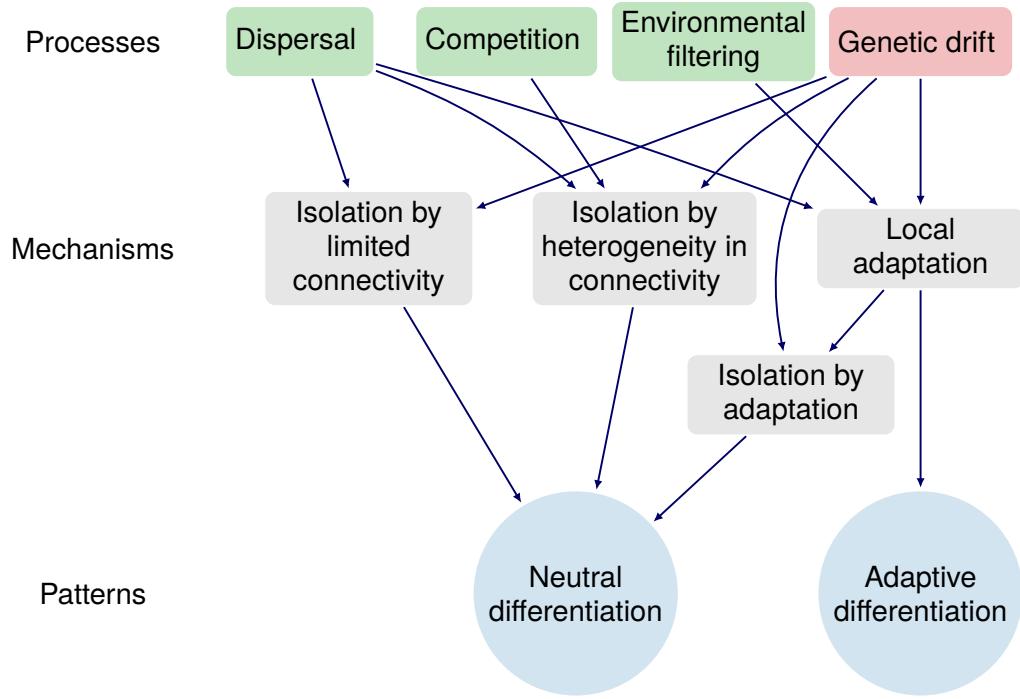


Fig. 2.1: Summary of the causal pathways involved in neutral and adaptive differentiation, disentangled in ???. Ecological processes are displayed in green boxes, evolutionary processes are displayed in red boxes.

dispersal. This results in habitat assortativity decreasing neutral differentiation for low dispersal intensity, and increasing neutral differentiation for high dispersal intensity. This complex feedback is essential to understand population differentiation in complex landscapes.

I provide a summary of the processes and resulting mechanisms shaping neutral and adaptive differentiation in Fig. 2.1. Overall, ??links processes to patterns, establishing a map of the causal pathways involved in the phenotypic differentiation of populations.

Linking economic patterns to eco-evolutionary processes

??provides a novel understanding of the role of eco-evolutionary processes in economic systems. Neoclassical economics and evolutionary economics seek to explain economic change with formal modelling, focusing on the relationships between economic variables such as output, employment and productivity [Boschma2005a]. In particular, evolutionary economics is concerned with explaining economic change by endogenous forces, such as interactions between firms and economic activities, and evolutionary processes acting upon them [Hodgson2019, Metcalfe2006]. In opposition to this approach, complexity economics [Hidalgo2021] seeks to predict

variations in national income by using fine-grained data of economic activity outputs and dimensionality reduction techniques [Mitchell], without assumptions on the underlying processes [Hidalgo]. While a current concern in evolutionary economics is to obtain an agreement between the mechanisms proposed and empirical observations [XXX], complexity economics seeks to understand the causal processes underlying the success of the dimensionality reduction technique [XXX].

By being agnostic to economic variables yet providing a causal link with the most probable processes underlying economic development, ??bridges evolutionary economics and economic complexity. Our approach relies on the simple observation that the dynamics of economic activities contains signatures from the many complex processes that underpin them. These signatures consist in peculiar temporal variations and couplings that can be conveniently modelled by biologically inspired population dynamic models, which use are additionally justified by deep analogies between economic activities and biological populations. Analogously to biological populations that are characterised by genes, economic activities are characterised by organizational routines [NelsonWinter], which experience evolutionary processes and define how they engage in ecological processes [NelsonWinter]. As a result, and similarly to the size of biological populations, the capital growth of economic activites is determined by the ensemble of organizational routines that characterise them [Boschma2005a]. Similarly to approaches used in ecology and evolution [Skeels], inverse modelling approaches can then be used to disentangle, from their signatures [Skeels] left on the dynamics of economic activities, the role of the eco-evolutionary processes acting upon them.

Specifically, ??explores the effect of eco-evolutionary processes, proposed in the evolutionary economics and geography economics literature, on the dynamics of economic activities at the national scale. In particular, ??seeks to test whether the dynamics of economic activities can be explained by different types of interdependencies, including positive [XXX]and negative interactions [XXX], spatial transfers [XXX], and transformations [XXX]. Using population dynamic models capturing the different interdependencies, ??provides empirical evidence that economic activities engage in positive interactions, and benefit from spatial transfers of knowledge and routines. Positive interactions may arise from a variety of processes proposed in the evolutionary economic literature, such as supply chains [Ozman2009, Saavedra2009a] and knowledge spillovers [Menon2015]. Its support implies that diversity promotes economic development [Hidalgo2018], as economic activities promote each other. The support for spatial transfers implies that, besides international patent laws [XXX], transfers of knowledge and organisational routines have a considerable on economic activities. Nevertheless, discrepancies in the strength of evidence obtained for spatial transfers across countries highlight that some countries

are overall more akin to spatial transfers than others, which may be explained by differences in cognitive, organizational, social, institutional or geographic proximities across countries [XXX]. I provide a summary of the mechanisms evidenced in economic systems in ???. Overall, ??bridges approaches in economics and biology to understand the mechanisms shaping the endogenous dynamics of economic systems.

2.1.2 Methodological advances in forward and inverse modelling

Advances in the modelling of realistic spatial and phenotypic structures

???? provide new tools to better understand mechanisms resulting from ecological and evolutionary processes, and their interplay, in the context of realistic population structures and phenotypic distributions. Evolutionary dynamics have been traditionally studied in the context of homogeneous or spatially extended populations [8]. For instance, [9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19] consider regular spatial structures, missing the effect of the complexity of spatial structures on population differentiation. Nevertheless, biological habitats differ in their connectivity [XXX], and economic entities are structured through complex networks [XXX]. The work of [8] and subsequent studies of "evolutionary dynamics on graphs" [XXX] show that this complexity affects the interplay between selection and drift. However, evolutionary dynamics on graphs does not consider eco-evolutionary feedbacks [Govaert2019a]. Thus far, models that include frequency dependence together with realistic, complex population structures were missing.

While a vast majority of the work on eco-evolutionary feedbacks has focused on the evolution of scalar phenotypes [Doebeli2011], in most organisms many phenotypic properties combine in complicated ways to determine ecological interactions, and hence frequency-dependent selection [Doebeli2014] [Doebeli2011] shows that the consideration of multiple traits with complex interactions relaxes the unrealistic conditions of strong frequency dependence required to generate diversity in one dimensional phenotype spaces, calling for a better understanding of evolutionary dynamics in high dimensional spaces. ?? demonstrates that the co-evolution of traits proved to have genuine consequences on differentiation, pointing towards the inclusion of multiple traits to understand the dynamics of ecological interactions. Other works on eco-evolutionary dynamics, including cancer cell evolution [XXX] and plankton dynamics [XXX], highlight the importance of considering multiple traits. The simulation of high dimensional models of trait leads to complications, since the numerical cost of traditional methods grows exponentially in the number of dimensions. Novel numerical methods have been proposed to simulate high dimensional models, but they could not handle non-local terms, which capture non-local interactions between microscopic agents.

?? develops a generic modelling framework to capture the effect of eco-evolutionary processes on biological populations structured in complex landscapes, and ?? provides tools to efficiently simulate the resulting high-dimensional models. The IBM presented in ?? involves the combination of graphs and highly dimensional phenotypic spaces, together with eco-evolutionary feedbacks, to model population structures. The modelling framework presented can readily be generalised, and the code associated to the numerical experiments in ?? include a Julia library, **Evoid.jl**, that implements a more general version of the model. As such, the modelling framework presented in ?? may be used to investigate other questions involving complex population structures and the co-evolution of characteristics.

Reproducing the discrete and stochastic nature of ecological and evolutionary processes [2], simulations of the IBM may not provide a general understanding of system investigated [XXX], and cannot be scaled to simulate large systems involving millions of individuals [XXX]. Nevertheless, the IBM proposed in ?? is mathematically tractable under simplifying assumptions, and can be efficiently simulated with a deterministic PDE approximation. Tractability allows to obtain analytical insights on how structural properties affect macroscopic population under simplifying assumptions (??). The PDE approximation, combined with the numerical methods presented in ??, further allow efficient simulations. I have implemented the numerical methods for simulating high dimensional models in the Julia library **HighDimPDE.jl** [**HighDimPDE**], a registered Julia package belonging to the SciML organisation [XXX]. The user interface respects standards from the SciML organisation, meaning that Julia users can easily adopt it. The package aims at hosting any solver algorithms that break down the curse of dimensionality, and is currently receiving contributions to implement the DeepBSDE scheme [20].

The combination of analytical insights and numerical simulations can help to elucidate the links between microscopic processes and macroscopic properties [Levin]. ???? provide novel analytical and numerical tools to link the interplay between eco-evolutionary processes and complex spatio-evolutionary structures to the resulting mechanisms and patterns.

Advances in inference methods for the investigation of eco-evolutionary processes

???? develop and test a novel inverse modelling method that allows to infer highly nonlinear dynamical processes from observation data. The method opens up new venues to further our understanding of the dynamics of biological and economic systems, and to provide predictions. The most developed inference methods in current use for inverse modelling are Bayesian inference methods with Markov Chain

Monte Carlo [XXX] and variational methods [XXX]. Bayesian inference methods require a large number of forward model integrations [21], and are highly affected by the number of model parameters [Csillary2010]. Variational methods require the model sensitivity to its parameters [XXX] and are prone to converge to local minima, especially with complex models [XXX]. Those central issues likely explain the very poor number of studies that have used inverse modelling to further our knowledge on eco-evolutionary dynamics [XXX].

?? presents a novel inference framework that can handle with high computational efficiency dynamic eco-evolutionary models. The framework is based on a simple variational method, but relies on recent development in ML, including AD [22] and state-of-the-art optimizers [23], and implements a learning strategy based on a mini-batch method, in order to circumvent the drawbacks of variational methods and adapt to the specificities of eco-evolutionary models. The use of AD simply eliminates the effort required to obtain the model sensitivity to its parameters, and the state-of-the-art optimizers, together with the mini-batch method, ensure the efficiency and reliability of the method in handling highly nonlinear models. By contrasting competing hypotheses embedded in alternative models, ??? provide concrete examples, both with synthetic and empirical data, that the ML framework can successfully elucidate eco-evolutionary mechanistic pathways. This work is part of an ongoing effort to blend ML and traditional models to gain scientific understanding and extrapolability [XXX]. Importantly, in contrast to physical systems such as ocean and atmospheric systems [XXX], invariant laws are yet to be formulated (if they exist). While ML is mostly used to improve model forecast skill in weather and climate modelling [XXX], the fields of Biology and Economic can gain scientific knowledge through model selection. Nevertheless, the proposed method is also relevant for improving the forecast of predictive eco-evolutionary model [24], and integrates the practical constraints of datasets including short time series with partial, noisy, shallow and independent observations [25].

The ML framework is implemented in the multi-purpose Julia package **MiniBatchInference.jl** [[MiniBatchInference](#)], readily available to the scientific community. **MiniBatchInference.jl** is built around the celebrated differential equation solver **DifferentialEquations.jl** and the deep learning library **Flux.jl**. As such, the use of **MiniBatchInference.jl** requires very limited efforts to any user familiar with those libraries.

Overall, the method proposed in ?? successfully blends ML methods with mechanistic ecosystem models to improve our gain scientific knowledge from observation data. Concret case examples in ??? show that it enables the testing of eco-evolutionary theories against data, advancing our understanding and the improvement of current

mechanistic models, with the potential to provide better forecasts of ecosystems states [24].

References

- [1] S. A. Levin. “Complex adaptive systems: Exploring the known, the unknown and the unknowable”. In: *Bulletin of the American Mathematical Society* 40.01 (2002), pp. 3–20. DOI: 10.1090/S0273-0979-02-00965-5.
- [2] N. Champagnat, R. Ferrière, and S. Méléard. “Unifying evolutionary dynamics: From individual stochastic processes to macroscopic models”. In: *Theoretical Population Biology* 69.3 (2006), pp. 297–321. DOI: 10.1016/j.tpb.2005.10.004.
- [3] S. Wright. “Isolation By Distance”. In: *Genetics* (1943).
- [4] B. H. McRae. “ISOLATION BY RESISTANCE”. In: *Evolution* 60.8 (2006), p. 1551. DOI: 10.1554/05-321.1.
- [5] B. H. McRae and P. Beier. “Circuit theory predicts gene flow in plant and animal populations”. In: *Proceedings of the National Academy of Sciences* 104.50 (2007), pp. 19885–19890. DOI: 10.1073/pnas.0706568104.
- [6] L. Orsini et al. “Drivers of population genetic differentiation in the wild: Isolation by dispersal limitation, isolation by adaptation and isolation by colonization”. In: *Molecular Ecology* 22.24 (2013), pp. 5983–5999. DOI: 10.1111/mec.12561.
- [7] T. J. Kawecki and D. Ebert. “Conceptual issues in local adaptation”. In: *Ecology Letters* 7.12 (2004), pp. 1225–1241. DOI: 10.1111/j.1461-0248.2004.00684.x.
- [8] E. Lieberman, C. Hauert, and M. A. Nowak. “Evolutionary dynamics on graphs”. In: *Nature* 433.7023 (2005), pp. 312–316. DOI: 10.1038/nature03204.
- [9] M. Slatkin. “GENE FLOW AND SELECTION IN A CLINE”. In: *Genetics* 75.4 (1973), pp. 733–756. DOI: 10.1093/genetics/75.4.733.
- [10] M. Slatkin. “Spatial patterns in the distributions of polygenic characters”. In: *Journal of Theoretical Biology* 70.2 (1978), pp. 213–228. DOI: 10.1016/0022-5193(78)90348-X.
- [11] M. Kirkpatrick and N. H. Barton. “Evolution of a Species’ Range”. In: *The American Naturalist* 150.1 (1997), pp. 1–23. DOI: 10.1086/286054.

- [12] J. Polechová and N. H. Barton. “Limits to adaptation along environmental gradients”. In: *Proceedings of the National Academy of Sciences of the United States of America* 112.20 (2015), pp. 6401–6406. DOI: 10.1073/pnas.1421515112.
- [13] J. Polechová. “Is the sky the limit? On the expansion threshold of a species’ range”. In: *PLoS Biology* 16.6 (2018), pp. 1–18. DOI: 10.1371/journal.pbio.2005372.
- [14] M. AndradeRestrepo, N. Champagnat, and R. Ferrière. “Local adaptation, dispersal evolution, and the spatial ecoevolutionary dynamics of invasion”. In: *Ecology Letters* 22.5 (2019). Ed. by V. Calcagno, pp. 767–777. DOI: 10.1111/ele.13234.
- [15] M. Doebeli and U. Dieckmann. “Speciation along environmental gradients”. In: *Nature* 421.6920 (2003), pp. 259–264. DOI: 10.1038/nature01274.
- [16] G. Meszéna, I. Czibula, and S. Geritz. “Adaptive Dynamics in a 2-Patch Environment: A Toy Model for Allopatric and Parapatric Speciation”. In: *Journal of Biological Systems* 05.02 (1997), pp. 265–284. DOI: 10.1142/S0218339097000175.
- [17] S. Yeaman and S. P. Otto. “Establishment and maintenance of adaptive genetic divergence under migration, selection, and drift”. In: *Evolution* 65.7 (2011), pp. 2123–2129. DOI: 10.1111/j.1558-5646.2011.01277.x.
- [18] F. Débarre, O. Ronce, and S. Gandon. “Quantifying the effects of migration and mutation on adaptation and demography in spatially heterogeneous environments”. In: *Journal of Evolutionary Biology* 26.6 (2013), pp. 1185–1202. DOI: 10.1111/jeb.12132.
- [19] S. Mirrahimi and S. Gandon. “Evolution of specialization in heterogeneous environments: equilibrium between selection, mutation and migration”. In: *Genetics* 214.2 (2020), pp. 479–491. DOI: 10.1534/genetics.119.302868.
- [20] J. Han, A. Jentzen, and W. E. “Solving high-dimensional partial differential equations using deep learning”. In: *Proc. Natl. Acad. Sci. USA* 115.34 (2018), pp. 8505–8510. DOI: 10.1073/pnas.1718942115.
- [21] T. Schneider et al. “Earth System Modeling 2.0: A Blueprint for Models That Learn From Observations and Targeted HighResolution Simulations”. In: *Geophysical Research Letters* 44.24 (2017), pp. 12,396–12,417. DOI: 10.1002/2017GL076101. arXiv: 1709.00037.
- [22] C. Rackauckas et al. “Universal Differential Equations for Scientific Machine Learning”. In: (2020). arXiv: 2001.04385.

- [23] D. P. Kingma and J. Ba. “Adam: A Method for Stochastic Optimization”. In: *arXiv:1412.6980* (2014), 15 pp. arXiv: 1412.6980.
- [24] M. C. Urban et al. “Improving the forecast for biodiversity under climate change”. In: *Science* 353.6304 (2016). DOI: 10.1126/science.aad8466.
- [25] M. Dornelas et al. “BioTIME: A database of biodiversity time series for the Anthropocene”. In: *Global Ecology and Biogeography* 27.7 (2018), pp. 760–786. DOI: 10.1111/geb.12729.

