

The wicked problem of biodiversity and ecosystem services in a changing world

Niamh Eastwood¹, William Stubbings¹, Mohamed Abdallah¹, Isabelle Durance², Jouni Paavola³, Martin Dallimer³, Jelena Pantel⁴, Samuel Johnson¹, Jiarui Zhou^{1,1}, James Brown⁵, Sami Ullah¹, Stefan Krause¹, David Hannah¹, Sarah Crawford⁶, Martin Widman¹, and Luisa Orsini¹

¹University of Birmingham

²Cardiff University Cardiff School of Biosciences

³University of Leeds

⁴The American University of Paris

⁵Lawrence Berkeley National Laboratory

⁶Goethe-Universitat Frankfurt am Main

September 28, 2020

Abstract

The Earth's biodiversity and the ecosystem services it provides are declining rapidly because of an increasing use of natural (finite) resources to meet human needs and environmental change affecting biodiversity dynamics. Truly trans-disciplinary solutions are needed for a sustainable development of ecosystem services. Collaboration and knowledge transfer across paleo biology, climatology, global ecology, evolution, biogeography, water sciences, computer science, statistics, and economics has to be mustered for scientists, resources managers and policy makers to counter ecosystem function collapse and ecosystem services loss. We propose a novel framework that closes the implementation gap, in which long-term dynamics of biodiversity, abiotic properties, and ecosystem functions are reconstructed using an unprecedented integration of biochemical and environmental fingerprinting of biological archives spanning centuries. The long-term dynamics obtained from this fingerprinting are then placed in a machine learning pipeline to identify cause-effect relations between environmental change and biodiversity dynamics. Predictive models are tested by hindcasting, and then used to accurately forecast the future of ecosystem services and their socio-economic impact under different climate change scenarios. The framework provides accessible tools to practitioners to translate cutting-edge research into practical solutions for environmental management and practice.

Background and rationale

Ecosystems are self-sustaining and self-regulating systems that provide ecosystem services (ES), including *provisioning* (food), *regulating* (e.g. climate), *supporting* (nutrient cycling, primary production), and *cultural* (e.g. aesthetic and recreational) services (Mace *et al.* 2012). A sustained delivery of services while maintaining the integrity of natural ecosystems is challenging because natural capital is finite and the impact of human intervention on ecosystems is uncertain and/or unknown across different spatial, temporal, and economic scales (DeFries & Nagendra 2017). Ecosystem management is complicated because complex and interdependent interactions among the components of an ecosystem create non-linear feedback, increasing the uncertainty of risks and unpredictable consequences (Fig. 1). Additionally, the ecological and economic values of these risks can diverge; limited resources, competing objectives and the need for economic profitability are at the core of the so-called 'wicked problem' - complex, open-ended and intractable - of ecosystem services (Jax *et al.* 2018).

In the last 50 years, 60% of ES have deteriorated or been overused (Mace *et al.* 2012). Rapid and severe biodiversity loss has been identified as its main cause (e.g. Cardinale *et al.* 2012). International governing bodies have set targets to preserve biodiversity and ES (e.g. the Convention of Biological Diversity 2012 and the Aichi Biodiversity Targets 2019). However, despite these efforts, the first Global Assessment of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) in 2020 found an accelerating decline in Earth’s biodiversity and associated services caused by an increasing use of natural (finite) resources to meet human needs (Ruckelshaus *et al.* 2020).

Mitigation interventions aimed at conserving and restoring natural capital have so far been inefficient and inadequate. This is in part because research on biodiversity and ES is constrained by disciplinary boundaries, whereas truly cross-disciplinary solutions are needed (Barnosky *et al.* 2016). Disciplinary approaches may neglect process interactions, result in research undertaken at inappropriate or disconnected scales, or use discipline-specific tools which are inadequate to address cross-disciplinary questions (e.g. Roux *et al.* 2017). Scientific knowledge has also rarely been translated into tools for stakeholders that address ecological, economic and social issues within the same framework (Ehrlich *et al.* 2012). Decision-making frameworks that enable the prioritization of intervention mechanisms based on objective evidence weighting ecological as well as socio-economic priorities simply do not exist. Relative to other forms of capital assets, natural capital is undervalued by governments, businesses and the public and recognized only upon its loss (Kinzig *et al.* 2011).

Novel conceptual and methodological approaches across disciplines must be combined, to provide an in-depth understanding of ecosystem complexities through a comprehensive screening of the spatio-temporal interdependencies of biodiversity, ecosystem functions, and ecosystem services. We propose a novel framework that can guide this necessary effort. Our proposed trans-disciplinary approach directly addresses issues associated with discipline constraints by: i) assessing the synergistic impact of multiple environmental factors on biodiversity unlike current measures which are often limited to dose-effect responses of single species to individual pollutants (Kanno 2016); ii) understanding the feedback between biotic and abiotic dynamics through space and time to quantify the cause-effect links between loss of biodiversity and of ecosystem functions which are typically not included in modelling efforts (Nicholson *et al.* 2019); and iii) understanding the role of evolution in driving biodiversity changes and affecting associated ecosystem functions and services (Baert *et al.* 2016). Recent work has highlighted that many conservation goals can be met only if the role of evolution is accounted for, yet the relationship between evolution and ecosystem services remains largely unexplored and omitted in model forecasts (Rudman *et al.* 2017); iv) providing a realistic valuation of costs and benefits of restoring ecosystem services, currently hindered by the lack of model forecasts and scenarios that incorporate food-web network (biotic interactions) properties and their feedback to ecosystem functions and services (Nicholson *et al.* 2019).

We focus on freshwater ecosystems and their services because these ecosystems are diverse, including drylands, wetlands, streams/rivers and ponds/lakes, and geographically widely distributed, ranging from the poles to the equator. Freshwater ecosystems are of high conservation value, delivering important ecosystem services (e.g. clean water, food provision and recreation) and being under increasing threat of destruction and degradation (Ruckelshaus *et al.* 2020). Freshwater ecosystems are also among the most impacted by biodiversity loss, with an 83% reduction in the global freshwater living planet index since 1970 (WWF 2018). Our framework enables the prioritization of interventions to restore natural capital and provide positive cultural outcomes, addressing the IPBES and the Convention on Biological Diversity priorities. Strategic prioritization of conservation actions is critical because socio-economic constraints prevent the restoration of all natural capital.

It all starts with biodiversity

People intuitively value biodiversity. However, different values are frequently assigned to different groups of organisms. Some see biodiversity only as a service, whereas others consider all organisms valuable, irrespective of any direct benefits derived from them (Mace *et al.* 2012). These two views are known as

the ‘ecosystem services perspective’ and the ‘conservation perspective’, respectively (Mace *et al.* 2012). In the former, a functional role is implicit but does not reflect the values that people hold for biodiversity beyond its functional role in ecosystem processes. In the latter, the role of biodiversity in underpinning ecosystem processes is ignored, and conservation efforts often focused on a subset of charismatic and/or threatened species (Smith *et al.* 2012). Neither of these perspectives captures the complexity and dynamics of biodiversity and its links to ES (Fig. 1).

In the early seventies May showed that the very existence of such biodiversity as we observe in nature is mathematically paradoxical (May 1972). Until then it had been widely assumed that the larger and more densely interconnected an ecosystem, the more stable it is – it was assumed that any fluctuations in the biomass of a given species would be less likely to cause runaway feedback processes leading to extinctions (Paine 1966). May showed that, if food webs were random graphs, then size and complexity (i.e. biodiversity and connectivity) should tend to destabilise communities. Hence, he argued that there must be some feature of food-web structure which counters this effect in nature.

Scientists have attempted to capture the essential features of food webs with various models (e.g. acyclic networks through the cascade model (Cohen & Newman 1985). More recently, the niche model (Williams & Martinez 2000) uses a one-dimensional axis to assign species to a position along a continuous interval of the axis, centred at a randomly chosen lower point, such that it consumes all species within the interval and none outside. Several refinements of the niche model have subsequently been proposed; yet they cannot solve May’s paradox, since, artificial food webs become more unstable with increasing size and complexity (i.e. density of interactions; Domínguez-García *et al.* 2019).

It has recently been shown that ecosystem models can become more stable with increasing size and complexity if the networks have a sufficient degree of “trophic coherence” (Johnson *et al.* 2014). Trophic coherence captures the extent to which the species in a food web (or the nodes in any directed network) fall into well-defined trophic levels. It can also be regarded as a measure of overall directionality: the degree to which biomass flows up the food web, from basal species to apex predators, rather than along less direct or circular paths. Trophic coherence underlies many other food-web properties (Klaire & Johnson 2017), including, crucially, the extent of feedback (Johnson & Jones 2017). It has, therefore, become apparent that to model food web complexity one needs to replicate the trophic coherence (e.g. Klaire & Johnson 2017).

Linking biodiversity to ecosystem function and service is challenging because of the inherent complexity of biodiversity (e.g. Durance *et al.* 2016). Growing evidence shows that biological diversity (the variation in genes, species, functional traits) is positively correlated to overall ecosystem function (resource capture, biomass production, decomposition, nutrient cycling) (Cardinale *et al.* 2012). Higher levels of biodiversity thus imply higher levels of ecosystem function and ecosystem services. This is convincing when considering provisioning services (e.g. provision of fish), or regulating services (decomposition, nutrient cycling), but is less obvious with services of a cultural nature. These services are defined by the beneficiary and they may thus differ across individuals, communities and generations (e.g. wars over water; Shiva 2002). However, because biodiversity can affect multifunctionality of ecosystems in non-additive ways compared to its effects on individual ecosystem processes (Perkins *et al.* 2015), preservation of increased volumes of biodiversity can be effective for maintaining ecosystem services. Ecologists have progressed to better understand what aspects of biodiversity - richness or evenness, in traits, genes, or phylogenies e.g. (Cadotte 2013); community structure and composition (Spaak *et al.* 2017) - contribute to ecosystem functions. De Laender (De Laender *et al.* 2016) recently described the importance of empirical manipulation of environmental drivers to better link environmental change, biodiversity shifts, and ecosystem function. Studies that span environmental change from community shifts to altered ecosystem functions are needed to gain a mechanistic understanding of current and future dynamics (Stockwell *et al.* 2020).

Establishing links between biodiversity and ES is further complicated by environmental factors affecting biodiversity dynamics, including ‘traditional’ (e.g., overexploitation, pollution, agriculture/land-use change) and ‘horizon’ (e.g., climate change) threats (Fig. 1) (Bonebrake *et al.* 2019). Chemical pollution, together with habitat loss, unsustainable use of resources, introduction of invasive species and climate change has

been identified as one of the main causes of biodiversity loss (Backhaus *et al.* 2012). This evidence combined with advances in analytical techniques, has triggered efforts to understand the detrimental impact of chemical pollutants on biodiversity beyond simple pollution scenarios (e.g. acid rain; Cardinale *et al.* 2012). Ultimately, the combined effect of multiple threats has to be considered, knowing that synergistic effects among environmental threats are responsible for at least 50 and up to 68% of species or community changes (Jackson *et al.* 2016). However, understanding biodiversity responses to multiple threats is challenging because they affect biotic interactions – e.g. insecticides impact on freshwater invertebrates increases nutrient pollution by (toxic) algal blooms by reducing grazing pressure (Alexander *et al.* 2013); ecological trade-offs – e.g. habitat modification can exacerbate the impacts of invasive predators (Doherty *et al.* 2015); and patterns of local adaptation - e.g. extinction most commonly occurs when threats are outside the evolutionary experience of species or outpace adaptation (Brook *et al.* 2008). Moreover, causes of biodiversity loss can operate on different spatial and temporal scales (Bonebrake *et al.* 2019) and are all context-dependent outcomes from processes operating over many years (Nogues-Bravo *et al.* 2018).

The promise of time traveling: continuous longitudinal data

Ecological processes, biodiversity dynamics, and environmental change occur over time. Therefore, longitudinal data are paramount to explain present-day patterns and to model cause-effect relations between biodiversity and environmental change (Baert *et al.* 2016). They are valuable for studying biodiversity dynamics bridging the gap between the timescale of empirical observations (at most 1 year) and the scales at which the underlying biological processes take place (at least multiple decades) (Balint *et al.* 2018). Long-term data also provide access to baseline observations predating major environmental impact to better understand long-term dynamics leading to current patterns (Nogues-Bravo *et al.* 2018). However, collecting these data requires commitments beyond research projects, and often beyond human lifespans. To compensate for scarce longitudinal data, researchers often use space-for-time-substitutions, studying the relationships between ecological variables at sites that are assumed to be at different stages of the same succession (e.g. clines) (Damgaard 2019). This is useful when the drivers that control biological turnover in time are the same as drivers that operate in space (Wogan & Wang 2018), and when the dynamics studied occur over short timeframes (Banet & Trexler 2013). However, it is problematic because it assumes static environments, requires that the observed spatial patterns are due to different realizations of the equilibrium properties of the studied ecological process, and that the temporal dynamics at the site level are ignored (Damgaard 2019).

A valuable source of longitudinal data are historical documents and paleo and archeo data (Lotze & Worm 2009). Studies in paleolimnology and **paleoecotoxicology** have used multiple lines of evidence, such as species assemblages, pollen records, nutrient loads, and historical contaminant profiles in dated sediment and ice cores, to improve environmental reconstructions that examine the impacts of land use changes, environmental stressors and other contaminants (Korosi *et al.* 2017). Yet, retrospective characterization of past effects of contaminants in aquatic ecosystems are poorly understood which may be explained by fossil remains analysis (**morphotaxonomy**) requiring specialist skills (e.g. light microscopy and taxonomy), being low throughput and limited to taxonomic groups with well-preserved remains in environmental matrices (e.g. pollen, exoskeletons, bones) (Gillson & Marchant 2014). The matrix-assisted laser desorption ionisation-time of flight-mass spectrometry (MALDI-TOF-MS) can identify specimens based on proteome fingerprinting, providing the advantage of high resolution mass spectra comparable across studies (Rossel & Arbizu 2018). For example, recent studies have applied the advanced **MALDI-TOF-MS** technique to study the impact of chemical pollutants on microbial biodiversity in environmental matrices such as seawater (Ashfaq *et al.* 2019) and settling dust (White *et al.* 2019). However, this technique relies on incomplete reference libraries to assign identities to investigate species based on percent match, and it can only be applied to living organisms, requiring high quality and quantity of input material (Rosse *et al.* 2019).

‘Ancient DNA (**aDNA** , Box 1)’ provides deeper reference baselines, enables the reconstruction of paleoecological and paleodemographic dynamics of particular species, and informs on species extinctions (Hambrecht 2017). Sedimentary ancient DNA (**sedaDNA** ; Box 1) has been used as an emerging tool to reconstruct

the past occurrence and trends of pollen, microbial and algal communities of lake systems for retrospective environmental assessment (Tse *et al.* 2018). However, these approaches only provide discrete measurements of taxonomic groups that leave persistent remains, such as bones, seeds, or scales (Hofman *et al.* 2015). Continuous biodiversity longitudinal data spanning evolutionary times coupled with environmental change records are needed to explicitly test hypotheses concerning the drivers of long-term biodiversity change. These data help understanding population continuity, migration, ecosystem structure and evolution (Orsini *et al.* 2013; Nogues-Bravo *et al.* 2018). Moreover, they enable one to disentangle the impact of extreme events (e.g. droughts, heat waves, toxic algal blooms) from the one of trends occurring on longer time scales (e.g. average temperature) (Jiguet *et al.* 2006). Continuous longitudinal data enable us to establish the cause-effect relationship between specific pollutants (e.g. pesticides) and biodiversity decline (Ashfaq *et al.* 2019).

Environmental DNA (**eDNA** ; Box 1) has been identified as a novel source of continuous longitudinal data in biodiversity (Balint *et al.* 2018), proving particularly advantageous when experimental manipulations are not feasible, early warning of ‘regime-shifts’ critical for ecosystem management have to be identified (Pace *et al.* 2015), and a broad taxonomic coverage is required to reconstruct community dynamics (Cristescu & Hebert 2018). Furthermore, **eDNA** is, to date, the only method that allows explicit analyses of ecological networks and species-species interactions through time (Domaizon *et al.* 2017); the detection of invasive, rare and endangered species in near real-time at affordable costs; and with non-specialist taxonomic skills (e.g. morphotaxonomy). Traditionally, **eDNA** required well-preserved DNA samples for an unbiased representation of the biodiversity, limiting the use of this technology to the recent past or undisturbed environments preserved in permafrost (Willerslev *et al.* 2003). However, recent advances in paleogenomics, which perform satisfactorily with low quality and low input DNA, have pushed back the temporal limits for the retrieval of so-called ‘dirty-DNA’ from different environmental matrices (**aeDNA** , Box 1), enabling analyses to be carried out at scales ranging from tens to hundreds and thousands of years (e.g. Pedersen *et al.* 2015).

Coupling continuous longitudinal biodiversity data with the continuous assessment of environmental change offers enormous potential to establish the cause-effect relationships between biodiversity loss and environmental variables. Possibly, the most complete long-term environmental data are climatic data. For most land areas, the temporal evolution of temperature since the beginning of the 20th century is well known from gridded station observations on spatial scales of a few hundred kilometres (IPCC 2013). Global gridded precipitation records with a spatial resolution of about 50km are available for land areas after 1901 (Meyer-Christoffer *et al.* 2013). Biodiversity impact studies require in principle local or regional climate data, which is not provided by these data. However, the last few decades are covered for many regions with resolutions of 1-20 km with gridded station records (e.g. Cornes *et al.* 2018) or merged station-satellite products (e.g. Navarro *et al.* 2019); for some study areas local station observations may also be available (e.g. Klein Tank *et al.* 2002). Additionally, local and large-scale temporal variability are often highly correlated, allowing the formulation of biodiversity impact models with large-scale climate variables as predictors. Together with climate change, chemical pollution is the main driver of biodiversity loss (Franklin & Raine 2019). With recent developments in high resolution mass spectrometry and metabolomics profiling on dated sediment samples, longitudinal reconstructions of pollutants’ dynamic are now possible (Chang *et al.* 2018). This approach, complemented by metabolomics of keystone species exposed to environmental pollutants enables investigations of the functional pathways disrupted by environmental pollution (e.g. pesticides; Moing *et al.* 2020).

Scale matters and analytical constraints

Biodiversity and ecosystem resilience are inherently related across spatial and temporal scales, with the heterogeneity of physical landscape properties controlling the connectivity and the distribution of, and interactions between, species. This includes strongly heterogeneous non-linear distributions of environmental properties and resulting non-linear ecosystem behaviour and dynamics, including hotspots and hot moments of abiotic ecosystem controls, such as point source pollution and extreme hydro meteorological events (Bernhardt *et al.* 2017). Observing and predicting spatially heterogeneous environmental properties and non-linear

ecosystem behaviour has been a challenge for operational and research monitoring networks and predictive models, often hampering the ability to adequately reflect the larger-scale (longer-term) consequences of hotspot (hot moment) disturbances or in fact, interventions (Krause *et al.* 2015). Low-cost sensor networks (LCSNs) (Mao *et al.* 2019), which largely focus on detecting and mitigating events with direct impact on humans (e.g. flooding an air pollution), have the potential for monitoring ecosystem structure and function, revealing the impact of ecosystem disturbance at different spatial and temporal scales.

While technological advances have lifted many of the limitations for continuous biochemical fingerprinting of environmental matrices, methodological advances in the analyses of these heterogeneous data within a single framework are often not applied because of disciplinary constraints. Data-driven modelling (DDM), in particular **multi-view learning** and **interpretable machine learning** (Box 2) have expanded analytical capabilities to interrogate complex, heterogeneous and non-replicated data, such as biological and environmental data (Willcock *et al.* 2018). These approaches identify and rank input variables based on causality, by establishing quantitative relationships between features and outputs (Schuwirth *et al.* 2019). For example, they can be used to establish cause-effect relations between environmental factors and species richness. **Representation learning** (RL) coupled with machine learning (Box 2) determines the representations needed for feature detection from raw data and can summarise the feature information in a higher-level abstraction, unveiling the underlying explanatory factors hidden in the observed low-level data (Bengio *et al.* 2013). Examples of RL applications are **random forest** and **deep learning** in hyperspectral imaging recognition for biodiversity and pollution monitoring (Vaglio Laurin *et al.* 2014), and in remotely sensed data applications for farm-specific management (Shyamal *et al.* 2020). An important next step is to link biodiversity and ecosystem function forecasts.

Some of the major challenges for biodiversity projections are the limited spatial resolution (approx. 100 km) of the state-of-the-art General Circulation Models (GCMs) that are used to generate future climate change projections (IPCC. 2018), and the difficulties in projecting regional climatic changes on the resolved spatial scales. Whilst the change in global mean temperature is dominated by the radiative effects of the greenhouse gases and can thus be simulated with relatively small uncertainties, regional changes in temperature and precipitation can carry substantial uncertainties due to biases in the simulated atmospheric circulation (e.g. Woollings 2010). These challenges can be partly alleviated by downscaling methods, which use either physically-based regional climate models (RCMs) with spatial resolutions of around 20km, or statistical models that use links between past large-scale and regional-scale climate variability to project future large-scale climate (for a review see Maraun & Widmann 2018).

A cross-disciplinary framework for resilient biodiversity and ecosystem services

The resilience of biodiversity and ecosystem services under the present and future climates is hindered by the lack of system-level approaches that provide a deep understanding of complex dynamics for prioritization of interventions to restore natural capital, revitalise stocks and flows within and between sites, and provide positive cultural outcomes, addressing the ambition of the Aichi Biodiversity Targets. In the following, we propose a novel cross-disciplinary framework that has the potential to revolutionize our understanding of past dynamics and the forecasts of biodiversity and ecosystem services under future climate scenarios (Fig. 2).

Step 1. Biochemical fingerprinting

A deep understanding of the context-dependent interactions between biodiversity and environmental change is needed. However, state-of-the-art measures of environmental change impact oversimplify natural processes and dynamics. Tracking biotic and abiotic changes through time and space studying biological archives (Orsini *et al.* 2013; Nogues-Bravo *et al.* 2018) enables us to lay the foundation for a knowledge-based framework that explains the relationship between system complexity and food-web energy and matter flows from pristine to degraded ecosystems. Choosing sedimentary archives that span at least a century allow

measures of biodiversity and environmental changes to be compared to a reference baseline predating the most significant period of human impact on climate and biodiversity, providing means to compare present-day changes to unperturbed baselines.

At the moment, regime shifts are measured in comparison to already shifted baselines, failing to elucidate the causes and processes causing the shifts. Recent developments in high throughput sequencing technologies applied to **toeDNA**, **aeDNA** and **sedDNA** (Box 1) extracted from specimens or remains in biological archives enable the capture of multiple trophic levels (e.g. prokaryotes and eukaryotes) across time and space, achieving a community-level understanding of biodiversity dynamics (Ruppert *et al.* 2019). Similarly, advances in gas and liquid chromatography coupled with high resolution mass spectrometry applied to different environmental matrices provide fingerprints of persistent pollutants over space and time (Hollender *et al.* 2019). Sedimentary archives can be dated using radioisotopes (Appleby 2001), enabling us to align community-level dynamics with environmental changes. Complementing biochemical observations through time and space with the analysis of Ecosystem Functions (e.g. biogeochemical functions measured as the accumulation rate of total organic, nitrogen and phosphorus) via bulk stoichiometry of sediments or other environmental matrices (phosphorous: P; nitrogen: N; carbon: C) helps elucidate long-term dynamics in productivity as influenced by nutrient availability and the relationships of stoichiometric ratios, productivity and biological attributes. The relationships among biodiversity attributes (measured via eDNA), environmental changes (measured via mass-spectrometry), and ecosystem functions (measured via bulk stoichiometry) will provide important insights into key functional changes in natural ecosystems. For example, past nutrient enrichment (eutrophication) linked to changes in biodiversity (species relative abundance and composition) provides insights into ecosystem shifts and tipping points of ecosystem stability. Longitudinal data collected with traditional methods (e.g. paleolimnology and morphotaxonomy) can be utilized to complement and confirm biochemical fingerprinting (Bennion *et al.* 2011). To be able to better disentangle temporal dynamics driven by stochasticity from the one driven by environmental change, similar undisturbed reference sites can be used as reference for temporal dynamics reconstructions.

Step 2. Establishing cause-effect relations between biotic and abiotic changes

Feedback between biotic and abiotic dynamics is ignored in current models that aim to quantify the cause-effect links between loss of biodiversity and of ecosystem functions with environmental change. Our framework establishes cause-effect relations among biodiversity attributes (measured via eDNA), environmental changes (measured via mass-spectrometry), and ecosystem functions (measured via bulk stoichiometry) by integrating multi-view learning, interpretable learning, and representation learning (Box 2; Fig. 2). This is achieved using Sparse Canonical Correlation Analysis (sCCA), a combination of feature selection and multi-view learning that enables the identification of the most relevant features in the multiple ‘views’, the linear combinations of which are maximally correlated in a common latent space (Tenenhaus *et al.* 2014). sCCA is extended as an interpretable multi-view learning algorithm by introducing co-responsive network analysis to identify both the inter- and intra- view associations (Tenenhaus *et al.* 2014). The algorithm can model the web of interacting ecological communities affected by the joint effect of environmental variables. Integrating sCCA with a Bayesian framework that accounts for trophic interactions (predation and herbivory), as well as mutualism, commensalism or parasitism, described by means of a multilayer network (Pilosof *et al.* 2017), allows empirical data to be complemented by models based on data obtained from similar ecosystems, providing a higher predictive power of causal associations. For example, the probability that species A preys on species B can be weighed against the mean trophic levels such species occupy in other food webs, and the trophic coherence of similar ecosystems. For other kinds of interactions other properties can be taken into account. For instance, the existence of symbiotic relationships might be a function of the expected nestedness of the mutualism layer of the network (Strona & Fattorini 2014).

Step 3. Forecasting biodiversity and ecosystem services under future climate scenarios

The feedback between biotic and abiotic interactions and species evolutionary potential are not accounted

for in forecasting models, leading to inaccurate estimates of biodiversity resilience (Urban *et al.* 2016). Continuous longitudinal data collected from biological archives are an unprecedented resource to guide predictions about future patterns and trends in biodiversity (Pace *et al.* 2015). Whereas past dynamics may differ from that caused by future environmental changes or threats, multidecadal information on food web architecture and eco-evolutionary feedbacks provide a powerful training sets for model forecasts to ‘learn’ from biodiversity assemblages and dynamics (Fordham *et al.* 2016). The integration of historical contamination trends with micro-evolutionary patterns in adaptation to pollutants’ exposure in a changing environment will help to further our understanding of risk trajectories associated with polluted sediments/aquatic ecosystems. Deep learning holds promises for predictive forecasting; we propose the use of Bayesian hierarchical mixed effects model and **iterative Random Forests** (iRF; Basu *et al.* 2018) for the identification of non-linear interactions between model features (e.g. biodiversity attributes, environmental and climatic variables and ecosystem biogeochemical functions) without the need for parametric assumptions. These non-parametric approaches combined with traditional parametric approaches (e.g. Oliver *et al.* 2015) can be used to derive consensus models using a generalized additive framework that generates predictions for changes in ecosystem functions driven by environmental change and mediated by changes in overall biotic interactions. For example, the “functional network” from the time series approach (De Vico Fallani *et al.* 2014) can be used to derive biotic interactions from biodiversity data obtained from the biological archives with information metrics, such as Granger causality, including: a) the effect of food-web trophic coherence on the scaling of stability with biodiversity; b) the importance of functional groups for ecosystem robustness; and c) the effects of eco-evolutionary feedbacks on ecosystem stability. Information provided by multidecadal dynamics during the transition from pristine to degraded state of natural ecosystems captured from the screening of biological archives, provides empirical longitudinal data against which iteratively optimized model forecasts can be validated (Fig. 2). Predictive models should be constructed to predict present-day dynamics from past changes; these predictions can then be validated by hindcasting, using empirical longitudinal data collected from biological archives as explained in (Nogues-Bravo *et al.* 2018). Once models have been optimized and validated against empirical data, past biodiversity-ecosystem dynamics that include measures of food web architecture, energy fluxes and eco-evolutionary feedbacks can be used in the Random Forests framework (iRF) together with temperature projections from state-of-the-art regional climate simulations (EURO-CORDEX; Jacob *et al.* 2014) as ‘training sets’ to predict the future of Ecosystem Functions under different climate scenarios. To reduce potential local model biases, the simulated climate change signal can be added to the observed local temperatures (delta change approach; Maraun & Widmann 2018). The performance of the models can be assessed using measures of local confidence and local feature importance, which is straightforward to do in the **iRF** framework. The iterative incorporation of new data into the models should be allowed for additional data collection directed specifically for better estimation of model parameters that contribute most strongly to predict uncertainty.

Step 4. Coupling ecological and economic value of ecosystem services

As ecosystems continue to degrade and are relied on by growing human populations for the services that they provide, tracking trends in values across ecosystems and contexts is vital for their conservation and restoration. A common approach is to assign monetary value to changes in ecosystem services which helps making direct comparisons with other costs and benefits in societal decision-making processes (Kahneman & Sugden 2005). The notion of monetary value can be thought of as a measure of a certain quality or level of importance (Schulz *et al.* 2017). For example, Costanza *et al.* (Costanza *et al.* 2014) estimated that the global value of ES loss between 1997-2011 was between \$4-20 trillion per annum depending on the estimation basis. Balmford *et al.* (Balmford *et al.* 2002) in turn demonstrated that the value of ecosystem services generated by intact ecosystems almost always exceeds the value of ecosystem services generated by modified or degraded ecosystems, and that an ambitious global programme of conservation could offer a cost-benefit ratio of 1:100. An outstanding challenge is the long lag between degradation and restoration, as well as between restoration and the realization of socio-economic benefits (Dallimer *et al.* 2015). For example, restored forested wetland in the Lower Mississippi River Valley had not regained the projected

level of denitrification for nitrate removal after 13 years (Ullah & Faulkner 2006). Another challenge is that values are often estimated over short, or unspecified timescales, which do not align with timescales of political and economic decision-making (Dallimer *et al.* 2020). Climate and environmental change can also undermine the resilience of ecosystems in the future, tipping the systems into new and unknown states; conventional economic valuation cannot handle these kinds of non-marginal changes (Turner *et al.* 2003).

The framework we propose can address these challenges by making use of both monetary and non-monetary arguments for conservation and restoration of ecosystems. An approach combining monetary and non-monetary arguments is needed because people’s perceptions of ecosystems and shared social values often differ from the suggestions of conventional economic models. Therefore, monetary valuation of ES as it has been traditionally practiced may not be appropriate for all facets of environmental goods such as non-use values (Nunes & van den Bergh 2001). Monetary value may not capture the full value of intact ecosystems in maintaining resilience against shocks and stresses and thus providing natural insurance: such insurance values, as well as option and quasi-option values of having the chance of future use of an ecosystem service have seldom been considered in valuation studies (Dallimer *et al.* 2020). Yet, they could provide further support for conserving and restoring ecosystems (Paavola & Primmer 2019). We suggest that handling and understanding the multi-dimensional value of ecosystems requires a combination of conventional and novel deliberative and participatory approaches (Kenter *et al.* 2015). The latter approaches still form a modest proportion of valuation studies, which reflects the rather recent recognition of the importance of incorporating the multi-dimensionality of value into ecosystem service assessments (Kenter *et al.* 2015). The consideration of these broader values also requires participatory decision making and valuation tools, such as multi-criteria analysis (Shreve & Kelman 2014).

By coupling ecological modelling with monetary and non-monetary valuation in our framework, we can quantify the costs and benefits of inaction as well as of proactive action to restore complexity in natural ecosystems and ES, alongside non-monetary arguments for doing so. Important insights are gained by using data collected from biological archives on specific **molecular operational taxonomic units**(MOTUs) e.g. fish communities by assessing whether replacing a native fish assemblage with a simplified system driven by e.g. an invasive species may have altered the ecosystem ecology and the livelihood of communities using fishing as a provisioning service. Coupled ecological-economic models can be used to assess the economic impact of biodiversity degradation on local communities under future scenarios, using IPCC projections up to 100 years in the future. For example, the economic loss associated with loss of provisioning and recreational services can be quantified.

Step 5. Interconnected systems: the importance of scale and dynamics of ecosystem controls

The developed framework has the potential of significantly advancing the state-of-the-art in understanding the feedback between biotic and abiotic interactions and in forecasting the resilience of ecosystems in different climate scenarios. This enables target biotic and abiotic interventions to achieve resilient ecosystems. However, ecosystems are not isolated units and the scale at which we analyse dynamics and patterns can significantly change decision making when it comes to prioritisation of interventions and mitigation strategies. In this respect, there is scope of utilizing the opportunities being released by the global sensor (network) revolution in environmental sensing technologies that we are experiencing globally (Mao *et al.* 2019). The resulting change in our ability to observe heterogeneous environmental conditions at unprecedented spatial and temporal scales and resolution causes step changes in our capacity for making reliable predictions of non-linear ecosystem behaviour. This provides new opportunities for adequately reflecting the impacts and implications of small-scale hotspots and hot moments of short duration onto large scale and long-term ecosystem stability and biodiversity as well as to assess the importance of long-term trends and large scale system changes onto regional and local ecosystem conditions (Krause *et al.* 2015). Ultimately, understanding the importance of small scale (short term) process dynamics on large scale (long-term) ecosystem behaviour will provide improved predictive capacity for assessing the likelihood of non-linear ecosystem responses to

multi-stressor interactions including tipping points and threshold behaviour (Dakos *et al.* 2019).

Implications for environmental practice and management

The proposed framework will enable to uncover the cause-effect links between abiotic and biotic changes, by quantifying dynamics (presence and relative abundance) of core **MOTUs** that are associated with ecosystem functions and abiotic changes. Using phylogenetic information, **MOTUs** can be placed into operational taxonomic units, allowing the analysis of compositional shifts and relative abundance of taxonomic units without a Linnaean name. This approach will enable the prioritization of mitigation interventions that preserve operational taxonomic units responsible for ecosystem stability. These interventions will favour the re-establishment of natural complexity and feedback, accelerating ecological restoration of natural ecosystems. This system-level approach using **MOTUs** requires the modernization of current environmental practice and a shift from bioindicator-based (morpho taxonomic of fossil remains (Gillson & Marchant 2014) or bioindicators, such as the Trophic Diatom Index (TDI; Kelly & Whitton 1995) to whole-system evidence-based interventions. The uptake of the novel framework will require that protocols for the high throughput biochemical fingerprinting of natural ecosystems are packaged into a *wetlab toolkit* that practitioners can implement in routine screening. This toolkit will likely complement traditional monitoring approaches e.g. light microscopy. Ideally, the initial screening of biochemical properties of a representative number of disturbed ecosystems as compared to pristine baselines and undisturbed ecosystems will help identifying more informative bioindicators. Similarly, the analytical framework for interpretation of past dynamics and the forecast of biodiversity and ES resilience should be made accessible to non-specialists. The accessibility of both the *wetlab* and the *predictive toolkit* can only be achieved if academics and practitioners work closely together to translate cutting-edge research into practical solutions for environmental practice and management.

Concluding remarks

Lack of understanding of the processes that underpin ecosystem services has often led to mismanagement with clear dis-benefits for the environment, the economy and human well-being. Approaches and tools constrained by discipline boundaries have been unable to capture process interactions and different spatio-temporal scales necessary to address the so-called wicked problem of biodiversity and ES. The divide between academia and practitioners as well as the lack of training for stakeholders and practitioners on state-of-the-art tools hinders the translation of science into practice, and results in outdated tools for environmental practice. Decision-making frameworks are needed to prioritize intervention mechanisms based on objective evidence weighting ecological as well as socio-economic priorities. The proposed framework, for the first time, combines machine learning algorithms with functional networks that account for the complex dynamics through time of both biotic and abiotic interactions to ‘train’ novel model forecasts on long-term bio-chemical and climate data. This novel predictive framework is applicable to all ecosystems and environmental conditions, because it finds signatures that recapitulate community dynamics (e.g. loss of MOTUs) driven by environmental change (e.g. warming, pollution) that alter ES (e.g. food provision). The framework provides knowledge and tools to manage natural capital sustainably through bottom-up interventions which benefit entire ecosystems and the services they provide. Mitigating the loss of MOTUs, magnifiers of ecosystem impact, and the targeted regulation of environmental factors with severe adverse effects on biodiversity improves ecosystem integrity and resilience. Resilience enables sustained delivery of ecosystem functions with a positive cascading effect on ES e.g., provisioning (e.g. food, clean water) and regulating services (e.g. climate mitigation, nutrient cycling).

Acknowledgments

We thank Chantal Jackson, School of Geography, Earth and Environmental Sciences at the University of Birmingham for the artwork in Figure 1 and William Scavone, Kestrel Studio, USA for the artwork in Figure

2. LO acknowledges NERC for support through the grant NE/N005716/1. The authors declare no conflict of interest.

References

- Alexander, A.C., Luis, A.T., Culp, J.M., Baird, D.J. & Cessna, A.J. (2013). Can nutrients mask community responses to insecticide mixtures? *Ecotoxicology* 22, 1085-1100.
- Appleby, P.G. (2001). *Chronostratigraphic techniques in recent sediments*. Kluwer Academic Publisher, Dordrecht, The Netherlands.
- Ashfaq, M.Y., Al-Ghouti, M.A., Qiblawey, H., Rodrigues, D.F., Hu, Y.D. & Zouari, N. (2019). Isolation, identification and biodiversity of antiscalant degrading seawater bacteria using MALDI-TOF-MS and multivariate analysis. *Science of the Total Environment*, 656, 910-920.
- Backhaus, T., Snape, J. & Lazorchak, J. (2012). The impact of chemical pollution on biodiversity and ecosystem services: the need for an improved understanding. *Integr Environ Assess Manag*, 8, 575-576.
- Baert, J.M., Janssen, C.R., Sabbe, K. & De Laender, F. (2016). Per capita interactions and stress tolerance drive stress-induced changes in biodiversity effects on ecosystem functions. *Nat Commun*, 7, 12486.
- Balint, M., Pfenninger, M., Grossart, H.P., Taberlet, P., Vellend, M., Leibold, M.A. *et al.* (2018). Environmental DNA Time Series in Ecology. *Trends Ecol Evol*, 33, 945-957.
- Balmford, A., Bruner, A., Cooper, P., Costanza, R., Farber, S., Green, R.E. *et al.* (2002). Economic reasons for conserving wild nature. *Science* 297, 950-953.
- Banet, A.I. & Trexler, J.C. (2013). Space-for-time substitution works in everglades ecological forecasting models. *PLoS One*, 8, e81025.
- Barabasi, A.L. & Oltvai, Z.N. (2004). Network biology: understanding the cell's functional organization. *Nature reviews genetics*, 5, 101-113.
- Barnosky, A.D., Ehrlich, P.R. & Hadly, E.A. (2016). Avoiding collapse: Grand challenges for science and society to solve by 2050. *Elementa: Science for the Anthropocene*, 4, 000094.
- Basu, S., Kumbier, K., Brown, J.B. & Yu, B. (2018). Iterative random forests to discover predictive and stable high-order interactions. *Proceedings of the National Academy of Sciences*, 115, 1943-1948.
- Bengio, Y., Courville, A. & Vincent, P. (2013). Representation Learning: A Review and New Perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35, 1798-1828.
- Bennion, H., Battarbee, R.W., Sayer, C.D., Simpson, G.L. & Davidson, T.A. (2011). Defining reference conditions and restoration targets for lake ecosystems using palaeolimnology: a synthesis. *Journal of Paleolimnology*, 45, 533-544.
- Bernhardt, E.S., Blaszczyk, J.R., Ficken, C.D., Fork, M.L., Kaiser, K.E. & Seybold, E.C. (2017). Control Points in Ecosystems: Moving Beyond the Hot Spot Hot Moment Concept. *Ecosystems* 20, 665-682.
- Bohmann, K., Evans, A., Gilbert, M.T., Carvalho, G.R., Creer, S., Knapp, M. *et al.* (2014). Environmental DNA for wildlife biology and biodiversity monitoring. *Trends Ecol Evol*, 29, 358-367.
- Bonebrake, T.C., Guo, F., Dingle, C., Baker, D.M., Kitching, R.L. & Ashton, L.A. (2019). Integrating Proximal and Horizon Threats to Biodiversity for Conservation. *Trends Ecol Evol*, 34, 781-788.
- Brook, B.W., Sodhi, N.S. & Bradshaw, C.J.A. (2008). Synergies among extinction drivers under global change. *Trends Ecol Evol*, 23, 453-460.

- Cadotte, M.W. (2013). Experimental evidence that evolutionarily diverse assemblages result in higher productivity. *Proceedings of the National Academy of Science USA* , 110, 8996–9000.
- Cardinale, B.J., Duffy, J.E., Gonzalez, A., Hooper, D.U., Perrings, C., Venail, P. *et al.* (2012). Biodiversity loss and its impact on humanity. *Nature* , 486, 59-67.
- CHANG, J., ZHANG, E.L., LIU, E.F., LIU, H.J. & YANG, X.Q. (2018). A 60-year historical record of polycyclic aromatic hydrocarbons (PAHs) pollution in lake sediment from Guangxi Province, Southern China. *Anthropocene* , 24, 51-60.
- Cios, K.J., Kurgan, L.A. & Reformat, M. (2007). Machine learning in the life sciences . *IEEE Engineering in Medicine and Biology Magazine* , 26, 14–16.
- Cohen, J.E. & Newman, C.M. (1985). A stochastic theory of community food webs I. Models and aggregated data. *Proceedings of the Royal society of London. Series B. Biological sciences* , 224, 421-448.
- Cornes, R., van der Schrier, G., van den Besselaar, E.J.M. & Jones, P.D. (2018). An Ensemble Version of the E-OBS Temperature and Precipitation Datasets. *Journal of Geophysical Research: Atmosphere* , 123.
- Costanza, R., R., d.G., Sutton, P., van der Ploeg, S., Anderson, S.J., Kubiszewski, I. *et al.* (2014). Changes in the global value of ecosystem services. *Global Environmental Change* , 26, 152–158.
- Cristescu, M.E. & Hebert, P.D.N. (2018). Uses and Misuses of Environmental DNA in Biodiversity Science and Conservation. *Annual Review of Ecology, Evolution, and Systematics* , 49, 209–230.
- Dakos, V., Matthews, B., Hendry, A.P., Levine, J.M., Loeuille, N., Norberg, J. *et al.* (2019). Ecosystem tipping points in an evolving world. *Nature Ecology and Evolution* , 3, 355–362.
- Dallimer, M., Davies, Z.G., Diaz-Porras, D.F., Irvine, K.N., Maltby, L., Warren, P.H. *et al.* (2015). Historical influences on the current provision of multiple ecosystem services. *Global Environmental Change* , 31, 307-317.
- Dallimer, M., Martin-Ortega, J., Rendon, O., Afionis, S., Bark, R., Gordon, L.J. *et al.* (2020). Taking stock of the empirical evidence on the insurance value of ecosystems. *Ecological Economics* , 167, 106451.
- Damgaard, C. (2019). A Critique of the Space-for-Time Substitution Practice in Community Ecology. *Trends Ecol Evol* , 34, 416-421.
- De Laender, F., Rohr, J.R., Ashauer, R., Baird, D.J., Berger, U., Eisenhauer, N. *et al.* (2016). Reintroducing environmental change drivers in biodiversity–ecosystem functioning research. *Trends in Ecology and Evolution* , 31, 905-915.
- De Vico Fallani, F., Richiardi, J., Chavez, M. & Achard, S. (2014). Graph analysis of functional brain networks: practical issues in translational neuroscience. *Philos Trans R Soc Lond B Biol Sci* , 369.
- DeFries, R. & Nagendra, H. (2017). Ecosystem management as a wicked problem. *Science* 356, 265-270
- Doherty , T.S., Dickman, C.R., Nimmo, D.G. & Ritchie, E.G. (2015). Multiple threats, or multiplying the threats? Interactions between invasive predators and other ecological disturbances. *Biological Conservation* , 190, 60-68.
- Domaizon, I., Winegardner, A., Capo, E., Gauthier, J. & Gregory-Eaves, I. (2017). DNA-based methods in paleolimnology: new opportunities for investigating long-term dynamics of lacustrine biodiversity. *Journal of Paleolimnology* 58, 1–21.
- Dominguez-Garcia, V., Dakos, V. & Kefi, S. (2019). Unveiling dimensions of stability in complex ecological networks. *Proceedings of the National Academy of Sciences* , 116, 25714-25720.
- Durance, I., Bruford, M.W., Chalmers, R., Chappell, N.A., Christie, M., Cosby, B.J. *et al.* (2016). The challenges of linking ecosystem services to biodiversity. . *Advances in Ecological Research* , 54, 87-134.

- Ehrlich, P.R., Kareiva, P.M. & Daily, G.C. (2012). Securing natural capital and expanding equity to rescale civilization. *Nature* , 486, 68-73.
- Fordham, D.A., Akcakaya, H.R., Alroy, J., Saltre, F., Wigley, T.M.L. & Brook, B.W. (2016). Predicting and mitigating future biodiversity loss using long-term ecological proxies. *Nature Climate Change* , 3086, 909-916.
- Franklin, E.L. & Raine, N.E. (2019). Moving beyond honeybee-centric pesticide risk assessments to protect all pollinators. *Nat Ecol Evol* , 3, 1373-1375.
- Garlapati, D., Charankumar, B., Ramu, K., Madeswaran, P. & Ramana Murthy, M.V. (2019). A review on the applications and recent advances in environmental DNA (eDNA) metagenomics. *Review in Environmental Science and Biotechnology* , 18, 389-411.
- Gillson, L. & Marchant, R. (2014). From myopia to clarity: sharpening the focus of ecosystem management through the lens of palaeoecology. *Trends Ecol Evol* , 29, 317-325.
- Hambrecht, G. (2017). Ancient DNA as a tool for navigating the Anthropocene *The Journal of Culture and Agriculture* , 39, 109-115.
- Hofman, C.A., Rick, T.C., Fleischer, R.C. & Maldonado, J.E. (2015). Conservation archaeogenomics: ancient DNA and biodiversity in the Anthropocene. *Trends Ecol Evol* , 30, 540-549.
- Hollenderl, J., van Bavel, B., Dulio, V., Farmen, E., Furtmann, K., Koschorreck, J. *et al.* (2109). High resolution mass spectrometry-based non-target screening can support regulatory environmental monitoring and chemicals management. *Environmental Science Europe* 31, 42.
- IPCC (2013). *Climate Change 2013: The Physical Science Basis* . Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
- IPCC. (2018). Global Warming of 1.5degC. In: *An IPCC Special Report on the impacts of global warming of 1.5degC above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty* (eds. Masson-Delmotte, V, Zhai, P, Portner, H-O, Roberts, D, Skea, J, Shukla, PR *et al.*).
- Jackson, M.C., Loewen, C.J.G., Vinebrooke, R.D. & Chimimba, C.T. (2016). Net effects of multiple stressors in freshwater ecosystems: a meta-analysis. *Global Change Biology* , 22, 180-189.
- Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O.B., Bouwer, L.M. *et al.* (2014). EURO-CORDEX: New high-resolution climate change projections for European impact research. *Regional Environmental Change* , 14 563-578.
- Jax, K., Furman, E., Saarikoski, H., Barton, D.N., Delbaere, B., Dick, J. *et al.* (2018). Handling a messy world: Lessons learned when trying to make the ecosystem services concept operational. *Ecosystem Services* , 29, 415-427.
- Jiguet, F., Julliard, R., Thomas, C.D., Dehorter, O., Newson, S.E. & Couvet, D. (2006). Thermal range predicts bird population resilience to extreme high temperatures. *Ecol Lett* , 9, 1321-1330.
- Johnson, S., Dominguez-Garcia, V., Donetti, L. & Munoz, M.A. (2014). Trophic coherence determines food-web stability. *Proceedings of the National Academy of Sciences* , 111, 17923-17928.
- Johnson, S. & Jones, N.S. (2017). Looplessness in networks is linked to trophic coherence. *Proceedings of the National Academy of Sciences* , 114, 5618-5623.
- Jordan, M.I. & Mitchell, T.M. (2015). Machine learning: Trends, perspectives, and prospects. *Science* , 349, 255-260.

- Kahneman, D. & Sugden, R. (2005). Experienced Utility as a Standard of Policy Evaluation. *Environmental and Resource Economics* , 32, 161-181.
- Kanno, J. (2016). Introduction to the concept of signal toxicity. *J Toxicol Sci* , 41, SP105-SP109.
- Kelly, M.G. & Whitton, B.A. (1995). The Trophic Diatom Index: a new index for monitoring eutrophication in rivers. *Journal of Applied Phycology* , 7, 433-444.
- Kenter, J.O., O'Brien, L., Hockley, N., Ravenscroft, N., Fazey, I., Irvine, K.N. *et al.* (2015). What are shared and social values of ecosystems? *Ecological Economics* , 111, 86-99.
- Kinzig, A.P., Perrings, C., Chapin, F.S., 3rd, Polasky, S., Smith, V.K., Tilman, D. *et al.* (2011). Sustainability. Paying for ecosystem services—promise and peril. *Science* , 334, 603-604.
- Klaise, J. & Johnson, S. (2017). The origin of motif families in food webs. *Scientific reports* , 7, 1-11.
- Klein Tank, A.M.G., Wijngaard, J.B., Konnen, G.P., Bohm, R., Demaree, G., Gocheva, A. *et al.* (2002). Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment. *International Journal of Climatology* , 22, 1441-1453.
- Korosi, J.B., Thienpont, J.R., Smol, J.P. & Blais, J.M. (2017). Paleo-ecotoxicology: What Can Lake Sediments Tell Us about Ecosystem Responses to Environmental Pollutants? *Environ Sci Technol* , 51, 9446-9457.
- Krause, S., Lewandowski, J., Dahm, C.N. & Tockner, K. (2015). Frontiers in real-time ecohydrology-A paradigm shift in understanding complex environmental systems. *Ecohydrology* , 8, 529-537.
- Kvist, S. (2013). Barcoding in the dark? A critical view of the sufficiency of zoological DNA barcoding databases and a plea for broader integration of taxonomic knowledge. *Mol Phylogenet Evol* , 69, 39-45.
- Li, Y., Yang, M. & Zhang, Z. (2019). A Survey of Multi-View Representation Learning. *IEEE Transactions on Knowledge and Data Engineering* , 31, 1863-1883.
- Lotze, H.K. & Worm, B. (2009). Historical baselines for large marine animals. *Trends Ecol Evol* , 24, 254-262.
- Mace, G.M., Norris, K. & Fitter, A.H. (2012). Biodiversity and ecosystem services: a multilayered relationship. *Trends Ecol Evol* , 27, 19-26.
- Mao, F., Khamis, K., Krause, S., Clark, J. & Hannah, D.M. (2019). Low-Cost Environmental Sensor Networks: Recent Advances and Future Directions. *Frontiers in Earth Science* , 7, 221.
- Maraun, D. & Widmann, M. (2018). *Statistical downscaling and bias correction in climate research* . Cambridge University Press.
- May, R.M. (1972). Will a large complex system be stable? *Nature* , 238, 413-414.
- Meyer-Christoffer, A., Rudolf, B., Schamm, K., Schneider, U. & Ziese, M. (2013). A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial (trend) analysis from 1901-present. *Earth System Science Data* , 5, 71.
- Moing, A., Allwood, J.W., Aharoni, A., Baker, J., Beale, M.H., Ben-Dor, S. *et al.* (2020). Comparative Metabolomics and Molecular Phylogenetics of Melon (Cucumis melo, Cucurbitaceae) Biodiversity. *Metabolites* , 10.
- Navarro, A., Garcia-Ortega, E., Merino, A., Sanchez, J.L., Kummerow, C. & Tapiador, F.J. (2019). Assessment of IMERG precipitation estimates over Europe. *Remote Sensing* , 11, 2470.
- Nicholson, E., Fulton, E.A., Brooks, T.M., Blanchard, R., Leadley, P., Metzger, J.P. *et al.* (2019). Scenarios and Models to Support Global Conservation Targets. *Trends Ecol Evol* , 34, 57-68.

- Nogues-Bravo, D., Rodriguez-Sanchez, F., Orsini, L., de Boer, E., Jansson, R., Morlon, H. *et al.* (2018). Cracking the Code of Biodiversity Responses to Past Climate Change. *Trends Ecol Evol* , 33, 765-776.
- Nunes, P.A.L.D. & van den Bergh, J.C.J.M. (2001). Testing socio-cultural valuation methods of ecosystem services to explain land use preferences. *Ecological Economics* , 39, 203-222.
- Oliver, T.H., Heard, M.S., Isaac, N.J., Roy, D.B., Procter, D., Eigenbrod, F. *et al.* (2015). Biodiversity and Resilience of Ecosystem Functions. *Trends Ecol Evol* , 30, 673-684.
- Orsini, L., Schwenk, K., De Meester, L., Colbourne, J.K., Pfrender, M.E. & Weider, L.J. (2013). The evolutionary time machine: using dormant propagules to forecast how populations can adapt to changing environments. *Trends Ecol Evol* , 28, 274-282.
- Paavola, J. & Primmer, E. (2019). Governing the provision of insurance value from ecosystems. *Ecological Economics* , 164, 106346.
- Pace, M.L., Carpenter, S.R. & Cole, J.J. (2015). With and without warning: managing ecosystems in a changing world. *The Ecological Society of America* , 13, 460-467.
- Paine, R.T. (1966). Food web complexity and species diversity. *The American Naturalist* , 100, 65-75.
- Pedersen, M.W., Overballe-Petersen, S., Ermini, L., Sarkissian, C.D., Haile, J., Hellstrom, M. *et al.* (2015). Ancient and modern environmental DNA. *Philos Trans R Soc Lond B Biol Sci* , 370, 20130383.
- Perkins, D.M., Bailey, R.A., Dossena, M., Gamfeldt, L., Reiss, J., Trimmer, M. *et al.* (2015). Higher biodiversity is required to sustain multiple ecosystem processes across temperature regimes. *Global change biology* , 21, 396-406.
- Pilosof, S., Porter, M.A., Pascual, M. & Kefi, S. (2017). The multilayer nature of ecological networks. *Nature Ecology & Evolution* , 1, 1-9.
- Rossel, S. & Arbizu, P.M. (2018). Effects of Sample Fixation on Specimen Identification in Biodiversity Assemblies Based on Proteomic Data (MALDI-TOF). *Frontiers in Marine Science* , 5, 1-13.
- Rossel, S., Khodami, S. & Arbizu, P.M. (2019). Comparison of Rapid Biodiversity Assessment of Meiobenthos Using MALDI-TOF MS and Metabarcoding. *Frontiers in Marine Science* , 6, 1-13.
- Roux, D.J., Nel, J.L., Cundill, G., O'Farrell, P. & Fabricius, C. (2017). Transdisciplinary research for systemic change: who to learn with, what to learn about and how to learn. *Sustainable Science* , 12.
- Ruckelshaus, M.H., Jackson, S.T., Mooney, H.A., Jacobs, K.L., Kassam, K.S., Arroyo, M.T.K. *et al.* (2020). The IPBES Global Assessment: Pathways to Action. *Trends Ecol Evol* , 35, 407-414.
- Rudman, S.M., Kreitzman, M., Chan, K.M.A. & Schluter, D. (2017). Ecosystem Services: Rapid Evolution and the Provision of Ecosystem Services. *Trends Ecol Evol* , 32, 403-415.
- Ruppert, K.M., Kline, R.J. & Md Rahman, S. (2019). Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. *Global Ecology and Conservation* , 17, e00547.
- Schulz, C., Martin-Ortega, J., Glenk, K. & Ioris, A.A.R. (2017). The Value Base of Water Governance: A Multi-Disciplinary Perspective. *Ecological Economics* , 131, 241-249.
- Schuwirth, N., Borgwardt, F., Domisch, S., Friedrichs, S., Kattwinkel, M., Kneis, D. *et al.* (2019). How to make ecological models useful for environmental management. *Ecological Modelling* , 411, 108784.
- Shiva, V. (2002). *Water wars: privatization, pollution and profit*. South End Press, Cambridge.
- Shreve, C.M. & Kelman, I. (2014). Does mitigation save? Reviewing cost-benefit analyses of disaster risk reduction. *International Journal of Disaster Risk Reduction* , 10, 213-235.

- Shyamal, S.V., P., V.K., Patil, V.C. & Jha , S.K. (2020). Remote sensing and machine learning for crop water stress determination in various crops: a critical review. *Precision Agriculture* , 21, 1121–1155.
- Smith, R.J., Verissimo, D., Isaac, N.J.B. & Jones, K.E. (2012). Identifying Cinderella species: uncovering mammals with conservation flagship appeal. *Conservation Letters* , 5, 205–212
- Spaak, J.W., Baert, J.M., Baird, D.J., Eisenhauer, N., Maltby, L., Pomati, F. *et al.* (2017). Shifts of community composition and population density substantially affect ecosystem function despite invariant richness. *Ecology Letters* , 20, 1315–1324.
- Stockwell, J.D., Doubek, J.P., Adrian, R., Anneville, O., Carey, C.C., Carvalho, L. *et al.* (2020). Storm impacts on phytoplankton community dynamics in lakes. *Global Change Biology* , 26, 2756–2784.
- Strona, G. & Fattorini, S. (2014). On the methods to assess significance in nestedness analyses. *Theory in Biosciences* , 133, 179–186.
- Tenenhaus, A., Philippe, C., Guillemot, V., Le Cao, K.A., Grill, J. & Frouin, V. (2014). Variable selection for generalized canonical correlation analysis. *Biostatistics* , 15, 569–583.
- Tse, T.J., Doig, L.E., Tang, S., Zhang, X., Sun, W., Wiseman, S.B. *et al.* (2018). Combining High-Throughput Sequencing of sedaDNA and Traditional Paleolimnological Techniques To Infer Historical Trends in Cyanobacterial Communities. *Environ Sci Technol* , 52, 6842–6853.
- Turner, R.K., Paavola, J., Farber, S., Cooper, P., Jessamy, V., Rosendo, S. *et al.* (2003). Valuing nature: lessons learnt and future research directions. *Ecological Economics* , 46, 493–510.
- Ullah, S. & Faulkner, S.P. (2006). Use of cotton gin trash to enhance denitrification in restored forested wetlands. *FOREST ECOLOGY AND MANAGEMENT* , 237, 557–563.
- Urban, M.C., Bocedi, G., Hendry, A.P., Mihoub, J.B., Pe'er, G., Singer, A. *et al.* (2016). Improving the forecast for biodiversity under climate change. *Science* , 353.
- Vaglio Laurin, G., Cheung-Wai Chan, J., Chen, Q., Lindsell, J.A., Coomes, D.A., Guerriero, L. *et al.* (2014). Biodiversity mapping in a tropical West African forest with airborne hyperspectral data. *PLoS One* , 9, e97910.
- White, J.K., Nielsen, J.L. & Madsen, A.M. (2019). Microbial species and biodiversity in settling dust within and between pig farms. *Environ Res* , 171, 558–567.
- Willcock, S., Martinez-Lopez, J., Hooftman, D.A.P., Bagstad, K.J., Balbi, S., Marzo, A. *et al.* (2018). Machine learning for ecosystem services. *Ecosystem Services* , 33, 165–174.
- Willerslev, E., Hansen, A.J., Binladen, J., Brand, T.B., Gilbert, M.T., Shapiro, B. *et al.* (2003). Diverse plant and animal genetic records from Holocene and Pleistocene sediments. *Science* , 300, 791–795.
- Williams, R.J. & Martinez, N.D. (2000). Simple rules yield complex food webs. *Nature* , 404, 180–183.
- Wogan, G.O.U. & Wang, I.J. (2018). The value of space-for-time substitution for studying fine-scale microevolutionary processes. *Ecography* 41, 1456–1468.
- Woollings, T. (2010). Dynamical influences on European climate: an uncertain future. *Philos Trans A Math Phys Eng Sci* , 368, 3733–3756.
- WWF (2018). Living Planet Report: aiming higher. (eds. Grooten, M & Almond, REA) Gland, Switzerland.
- Xu, C. & Jackson, S.A. (2019). Machine learning and complex biological data. *Genome Biology* , 20, 76.

Box 1

The DNA maze: aDNA, eDNA, aeDNA, and sedaDNA

Traditional biodiversity monitoring techniques are problematic due to difficulties associated with correct identification of (cryptic) species or juvenile life stages, limited specialist taxonomic expertise, non-standardized sampling, the invasive nature of some survey techniques and the low throughput outcome. In the last few years, DNA-based methods for large-scale biodiversity monitoring have gained momentum (Balint *et al.* 2018). DNA extracted from environmental matrices is referred to as ‘Environmental DNA’ or **eDNA**; this term emerged at the beginning of the 2000s (Cristescu & Hebert 2018). eDNA has the potential to provide continuous temporal data as opposed to ancient DNA (**aDNA**), which is extracted from organisms that leave behind visible traces in archaeological sites (bones, seeds, scales) (Hofman *et al.* 2015), and palaeoenvironmental DNA (**aeDNA**) which incorporates soft-bodied and macrofossil organisms (Gillson & Marchant 2014). Sedimentary ancient DNA (**sedaDNA**) is emerging as a popular tool to reconstruct the past occurrence of microbial communities of inland waters for retrospective environmental assessment (Tse *et al.* 2018).

DNA extracted from environmental matrices and macrofossils combined with genomics and high throughput sequencing technologies informs conservation biology by estimating human impact on biodiversity (Hofman *et al.* 2015), invasion biology by identifying timing and severity of alien species invasion (Ruppert *et al.* 2019) and biodiversity science by enabling the census of species/taxa on a global scale in real time (Garlapati *et al.* 2019). In addition, eDNA provides continuous temporal data that are key to identifying temporal baselines and to assess changes in biodiversity over time. Quantitative identification of single species can be achieved using targeted quantitative PCR (**qPCR**), an accurate quantitative method applied to eDNA, aDNA, aeDNA and *sed* aDNA. It is a low-costs/high accuracy approach, suitable for single species identification. However, it requires knowledge of the genetics of the species of interest for detection probe design. Semi-quantitative, whole community biodiversity data can be obtained in a high throughput fashion by using multiple markers (metabarcoding) or total eDNA sequencing (metagenomics). As eDNA-based approaches move from single-marker to multi-marker approaches, eDNA becomes the method of choice to reconstruct dynamics of past and present communities without the need of specialist taxonomic skills and of organisms remains (Gillson & Marchant 2014). However, the poor representation of some ecosystems, taxa and genes in reference databases used for taxonomic assignment can lead to unassigned or incorrectly assigned sequencing reads (Kvist 2013). Although technical challenges associated with eDNA degradation and analytical limitations associated with incomplete reference databases exist, multiple markers eDNA applied to longitudinal samples has enormous potential, including an improved ability to explore ecosystem-level processes and community diversity, identifying tipping points of ecosystem stability and novel opportunities of detecting rare or difficult-to-sample taxa. Longitudinal studies on multidecadal or centennial time spans enable the unprecedented survey of entire ecosystems to predict spatial and temporal biodiversity patterns (Bohmann *et al.* 2014).

Box 2

Machine learning in biology

With technological advances, biochemical data generation is no longer limiting, whereas analytical approaches have become the bottleneck in current biology research (Xu & Jackson 2019). Large volumes of ‘imperfect’ empirical data (e.g. low replication, missing values, and heterogeneous samples) and the need to understand the mechanisms underpinning biological dynamics pose significant challenges to traditional statistics approaches. Conversely, **ML** holds unprecedented promises to extract information from complex and diverse data, because it manages to “learn” from input features with limited parametric assumptions on the data (Jordan & Mitchell 2015). Primary **ML** approaches include **unsupervised learning** e.g., feature extraction and clustering, and **supervised learning** e.g., classification and regression; both have been applied in biology for biomarker discovery, genomic feature prediction, biomonitoring via DNA metabarcoding, and population genomics (Cios *et al.* 2007). However, these approaches suffer from the “curse of dimensionality”, that is they require large sample sizes.

Representation learning (RL) or feature learning (FL) ably handles small sample sizes and many features, disregarding variables that contribute minimally to explaining features; extracting high-order associations, and identifying informative abstraction (i.e. “representation”) of the predictive features (Bengio *et al.* 2013). **Multiview learning** enables the use of multiple data sources to support the model training, contributing to knowledge needed to fully explain the different aspects of a complex biological system. It helps to identify consensus and complementary information originating from different data sources to produce predictions (Li *et al.* 2019). Interpretability is a common issue in **ML** due to the black-box nature of commonly used models, which hinders biological discoveries. Recently, increasing efforts have been devoted to the development of **interpretable ML**, which supports data-driven model training when the information from empirical samples is incomplete, being apt for knowledge integration. Interpretability can be achieved by providing an explanation for the predictive model. For instance, **random forest (RF)** is powerful to predict the relationship between feature and response outputs without relying on the assumption of data distribution. **iRF** improves the interpretability of conventional RF by creating a reduced-order representation from a subset of decisive features, which enables not only more effective intervention approaches but also an insight into the underlying mechanism (Basu *et al.* 2018). **ML** combined with Network Biology, which “understands and models in quantifiable terms the topological and dynamical properties of the various networks that control the behaviour” of biological processes, enables accurate interpretation of extant and future ecosystem dynamics (Barabasi & Oltvai 2004).

Glossary

Paleoecotoxicology applies environmental toxicology and chemistry to the interpretation of the fossil record.

Paleolimnology is the science that uses the physical, chemical, and biological information preserved in freshwater sedimentary archives to reconstruct past environmental conditions.

Morphotaxonomy is a taxonomic classification of species based on whole or partial body parts.

aDNA or ancient DNA is DNA isolated from ancient specimens, including bones, pollen, and seeds.

aeDNA or palaeoenvironmental DNA is DNA extracted from paleo environmental matrices.

sedDNA is DNA extracted from sediment matrices, both from non-obvious biological source material and from viable resting stages.

eDNA or environmental DNA is DNA extracted from environmental matrices, such as soil, sediment, water, ice and aerosol without any obvious signs of biological source material.

qPCR or quantitative polymerase chain reaction is a laboratory technique used for measuring the amplification of a targeted DNA molecule real-time during PCR.

Metabarcoding is high throughput sequencing of PCR-amplified taxonomic marker genes.

A **MOTU** or Molecular Operational Taxonomic Unit is a cluster of sequence variants obtained from marker gene sequences. MOTUs identify taxa through sequence identity, which does not necessarily correspond to Linnaean species identification.

MS or mass spectrometry is an analytical technique that measures the mass-to-charge ratio of ions. It is used to calculate the exact molecular weight of molecules in a given sample

MALDI-TOF-MS or matrix assisted laser desorption ionization-time of flight mass spectrometry uses a laser energy absorbing matrix to create ions from large molecules for the analysis of bio and organic molecules.

DL or deep learning is a ML method based on artificial neural networks with representation learning. Learning can be supervised, semi-supervised or unsupervised.

Feature learning (FL) or representation learning (RL) is a set of techniques that allows a system to automatically discover the representations needed for feature detection or classification from raw data.

ML or Machine learning is a subset of artificial intelligence and focuses on the development of algorithms that can access data and use it to learn for themselves

Interpretable ML is a ML algorithm that produces human-understandable outcomes.

iRF or Iterative Random Forest is an algorithm which is computationally efficient in searching for interactions of unknown form and order in high dimensional data.

Multi-view learning is an emerging direction in machine learning which considers learning with multiple views to improve the generalization performance.

RF or Random Forest is a machine learning tool that fits many classification or regression trees to a data set (using many bootstrap samples from the data and random variable selection) then combines the predictions from all the trees using model averaging.

sCCA or Sparse Canonical Correlation Analysis is a multivariate statistical method which describes correlations among subsets of variables from multiple data sets.

Figure legends

Figure 1 . Biodiversity and ecosystem services.

Biodiversity dynamics is affected by environmental change, including traditional (e.g. pollution, land use) and horizon (e.g. climate change) threats. The feedback between biodiversity and environmental change can alter ecosystem health and resilience, resulting in loss of functions and services, including - *provisioning* (food), *regulating* (e.g. climate), *supporting* (nutrient cycling, primary production), and *cultural* (aesthetic, social, recreational) services. Loss of ecosystem services has an adverse impact on society and the economy and can trigger interventions to restore biodiversity, which is the primary cause of loss of services. Interventions to re-establish biodiversity can be, however, hindered by environmental change, with adverse effects on human and ecosystem health.

Figure 2. Novel trans-disciplinary framework

Biodiversity, ecosystem function (EF), ecosystem services (ES) and environmental change (EC) are measured via fingerprinting of biological archives. B-EF and B-EC associations are established via explainable machine learning combined with functional networks and validated experimentally. Model forecasts that ‘learn’ from past dynamics and feedback are tested iteratively against real data (hindcasting) and refined to predict the future of B, ES and their economic value under different climate scenarios.

