Network-based metrics of resilience and ecological memory in lake ecosystems

David I. Armstrong McKay*[1,2,3], James G. Dyke[1,4], C. Patrick Doncaster[5], John A. Dearing[1], Rong Wang[6]

[1] *Geography and Environmental Science, University of Southampton, Southampton, UK, SO17 1BJ (work started here)*

[2] *Stockholm Resilience Centre, Stockholm University, SE-10691 Stockholm, Sweden (current address DIAM)*

[3] *Bolin Centre for Climate Research, Stockholm University, SE-10691 Stockholm, Sweden*

[4] *Global Systems Institute, College of Life and Environmental Sciences, University of Exeter, Exeter, UK (current address JGD)*

[5] *School of Biological Sciences, University of Southampton, Southampton, UK, SO17 1BJ*

[6] *State Key Laboratory of Lake Science and Environment, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, China, 210008*

***david.armstrongmckay@su.se**

Paper to be submitted to: PLOS Computational Biology

Abstract

Some ecosystems undergo abrupt transitions to a new regime after passing a tipping point in an exogenous stressor, for example lakes shifting from a clear to turbid 'eutrophic' state in response to nutrient-enrichment. Metrics-based resilience indicators have been developed as early warning signals of these shifts but have not always proved reliable indicators. Alternative approaches focus on changes in the composition and structure of an ecosystem, which can require long-term food-web observations that are typically beyond the scope of monitoring. Here we prototype a network-based algorithm for estimating ecosystem resilience, which reconstructs past ecological networks solely from palaeoecological abundance data. Resilience is estimated using local stability analysis, and eco-net energy: a neural network-based proxy for 'ecological memory'. We test the algorithm on modelled (PCLake+) and empirical (lake Erhai) data. The metrics identify increasing diatom community instability during eutrophication in both cases, with eco-net energy revealing complex eco-memory dynamics. The concept of ecological memory opens a new dimension for understanding ecosystem resilience and regime shifts, with eco-memory potentially increasing ecosystem resilience by allowing past memorised eco-network states to be recovered after disruptions.

Keywords: Lake Eutrophication; Early Warning Signals; Resilience; Palaeoecology; Neural Networks; Regime Shifts; Tipping Points

Introduction

The potential for stressed ecosystems to tip abruptly into a new regime has led to a proliferation of metrics attempting to quantify ecosystem resilience [1–10]. Lake eutrophication is a well-studied example of such a regime shift, and often exhibits evidence of alternative stable states and hysteresis [11–13]. Eutrophication occurs when increasing nutrient loading triggers positive feedbacks involving lake anoxia and light reduction within the ecosystem, driving a rapid shift from clear to turbid water conditions [14,15]. Recovery to clear conditions often requires a reduction in nutrient levels far below the original threshold, indicating hysteresis [13], although this is not a characteristic of all lake systems [11]. Prior to tipping, the ecosystem experiences declining resilience, defined here as the weakening of negative (stabilising) feedbacks relative to positive (amplifying) feedbacks, resulting in greater sensitivity to small shocks and longer recovery times (i.e. 'engineering' resilience) [16,17]. Attempts to develop ecosystem resilience indicators in order to understand and monitor the drivers of regime shifts broadly fall into three approaches: time-series metrics, compositional analysis, and network-based analysis.

Monitoring ecosystem resilience

Many resilience indicators test for 'critical slowing down': a slowing recovery rate from perturbations and increasing variability, which have been detected in various environmental time-series [4–10]. These time-series metrics have had inconsistent success, however, when applied as early-warning signals (EWS) of tipping points in freshwater [18] and other ecosystems [19–22]. A key methodological issue is the need for prior knowledge of one response variable that captures the dynamics of the whole system when analysed. Other limitations include possible false positives or negatives (where EWS indicate an impending transition which never occurs, or is absent prior to a known transition), and sensitivity to subjective time-series analysis parameter choices [7,23,24]. These problems are partially

resolved by using multiple sensitivity-tested metrics as generic resilience indicators rather than as EWS of specific critical transitions [16,25]. Palaeorecords present additional problems, with variable temporal resolution from changing sediment accumulation rates, compaction, or dating uncertainty making robust time-series analysis challenging [26,27].

As an alternative to metrics-based EWS, analysis of ecosystem composition seeks to detect changes in community functional dynamics without having to select or understand the full trophic ecology of any one species. To this end, Doncaster et al. [28] quantified the compositional disorder of diatoms and chironomids from sediment cores in three Chinese lakes, where low disorder signifies highly-nested sequential compositions. Several decades prior to critical transitions in these lakes, the correlation of disorder with biodiversity becomes negative and strengthens towards the tipping point. Theory and simulations [28] suggest that nutrient loading had shifted the competitive balance from 'weedy' (weakly competitive, fast-replicating) towards 'keystone' (strongly competitive, slow-replicating) species. This correlative approach avoids issues of variable temporal resolution, but the link with ecosystem resilience is indirect and the competition dynamics remain hypothetical. Another recent composition-based method identified negative skewness in nodal degree (connectivity) within a network of interacting diatom species, in response to increasing nutrient input into Chinese lake ecosystems [29]. This trend towards strong interactions is compatible with keystone dominance rising with exogenous stress.

Network-based methods represent another whole-ecosystem approach, which allows the possibility of performing local stability analysis on reconstructions of lake food-webs. Kuiper et al. [30] showed that food-web data in the form of material flux descriptions can be used to reconstruct interaction strengths between different species, which correspond to the interaction coefficients of a Lotka-Volterra ecosystem model (Figure 1). The matrix of these coefficients is equivalent to a Jacobian matrix of the Lotka-Volterra equations when considered as a dynamical system, and so methods for determining the local stability of

dynamical systems can also be applied to this interaction matrix. These systems are considered locally stable if the real part of the Jacobian's eigenvalues remain negative (representing stabilising net-negative feedbacks) [30–32]. A critical transition occurs when an eigenvalue crosses the zero line (representing the emergence of destabilising net-positive feedbacks). Given food-web measurements, the local stability of the food web can be estimated from the intraspecific interaction strengths, and is closely related to the Jacobian's dominant eigenvalue, λ_d (the largest absolute eigenvalue of the Jacobian matrix) [30,33–35]. The value of λ_d increases with increasing post-perturbation recovery times as a result of the weakening of negative relative to positive feedbacks. It thus functions as a proxy for weakening ecosystem resilience. An application of this approach to the PCLake model system enabled tracking of destabilising food-web reorganisations prior to eutrophication [30].

Ecological network inference

Although the local network stability method depends less on temporal resolution than time-series analysis [30] and avoids reliance on choosing a single variable, it requires gathering detailed food-web data at regular intervals during eutrophication, a laborious [36] and often unmanageable task for real-world lakes. Many lakes nevertheless have palaeoecological records of species abundances through time, obtained from sediment cores. If past ecosystem interactions can be reconstructed from these data, it becomes possible to estimate past changes in ecosystem resilience. Network inference is a novel technique which allows ecological networks (eco-nets) to be reconstructed from abundance data [37–40]. This technique for characterising interspecific interactions comes from microbial metagenomics, where it provides an alternative to the unreliable proxy of assuming that correlations between the relative abundances of different species indicate causative interactions between them [39]. Network inference instead uses multilinear regression to infer the interaction coefficients of a discrete-time Lotka-Volterra ecosystem model from abundance data (detailed in Materials and

Methods), with nodes in the inferred network representing individual species or functional groups and edges their interactions (Figure 1). Network inference thus allows the past structure of ecological communities to be reconstructed from palaeoecological abundance data, and is less sensitive to irregular temporal resolution than time-series analysis [38]. Ecosystem resilience can then be explored by performing local stability analysis on the inferred interaction matrix with a rolling temporal window, in order to track long-term changes in λ_d.

Ecological memory and eco-net energy

In this study we also develop and apply a novel neural network-based method of resilience analysis. Analyses of Lotka-Volterra systems demonstrate how an ecosystem can retain a distributed 'memory' of past states as a result of a process akin to unsupervised Hebbian learning in a type of neural network known as a Hopfield network [41,42]. Neural networks are computing systems loosely based on brains, with nodes (neurones) connected by edges (axons/synapses) used to process information. Neural networks are trained to recognise and process particular input data (the training input), with 'learning' indicating adaptation of the network's connections in response to the training, and the training input then stored as a distributed 'memory' as the network's structural adaptations (Figure 2a). In many networks this process is externally guided to reach desired results ('supervised learning') and is not applicable in natural ecosystems, but unsupervised 'Hebbian' learning can also lead to memory spontaneously emerging when some neural networks are given training input. In Hopfield networks for example, correlation in activity between two nodes leads to an increase in the connection strength between them – a process described colloquially as "*neurons that fire together, wire together*" (Hebb's Rule [43–46]). Over time this allows the emergence of distributive associative memory of training input that can be recovered when given degraded input (Figure 2b).

Power et al. [41] proposed that eco-nets experience similar learning dynamics, with species (nodes) that frequently co-occur developing stronger interactions (edges), allowing emergence of a distributed associative "ecological memory" (eco-memory) of past environmental forcing (training input) that acts as a stable attractor (Figure 2). This process has been simulated in a simple Lotka-Volterra model in which species are subject to natural selection by applying environmental forcing in the form of altered individual carrying capacities (equivalent to training input), with this training input subsequently recoverable even when the model is reset with new abundances [41]. This ecological memory can increase the ecosystem's resilience by acting as a stable attractor towards which the ecosystem recovers after perturbation. These findings suggest that ecosystems can learn from environmental forcing and become more resilient as a result of network effects [42], but this process has not yet been tested in more realistic process-based models or empirical data.

In this paper, we quantify the strength of ecological memory from abundance data as a metric of resilience, rather than recovery of training inputs *per se*. Although memory strength has no direct metric, one can calculate a Hopfield network's 'energy'. Training a Hopfield network results in a stable minimum forming in its associated energy function, which represents the memorised state. Once trained, this minimum acts as an attractor when the network is given new input (Figure 3a). Because network energy decreases with learning, and is minimised at locally stable points, it indicates how much the network has learnt about its past training. We can calculate eco-net energy, E_N, by treating the interaction matrix reconstructed by our network-inference algorithm as the weight matrix [41] of a continuous Hopfield network [43–46] (see Materials and Methods), and use it to indicate the proximity of the eco-net to the bottom of a memorised attractor as well as the depth of the attractor (Figure 3, right). We would expect low E_N for eco-nets that have 'learnt' from stable environmental conditions or have shifted to a more stable memory attractor (Figure 3a), and higher E_N when ecosystem destabilisation perturbs the eco-net away from its learned state or shifts it to a less

stable memory attractor (Figure 3b). Unlike a conventional Hopfield network though, eco-nets can also continue to learn from new conditions. Thus, a novel stable ecosystem state can form a new memory ('relearning') and energy minimum, potentially leading to lower E_N after some initial destabilisation (Figure 3c).

Study Aims and Design

In this paper we prototype and test an algorithm in R [47] for reconstructing ecological networks (eco-nets) from palaeoecological data and determining past changes in resilience using network-based metrics. We first use network inference to reconstruct the ecosystem's Lotka-Volterra interaction matrix from abundance data on a rolling temporal window, before exploring the ecosystem's changing resilience using two different approaches. Firstly, we perform local stability analysis on the inferred interaction matrices in order to track changes in λ_d. We test this analysis on two different datasets of eutrophication-induced regime shifts: whole ecosystem output from a commonly-used lake ecosystem model (PCLake+) for a hypothetical lake, and empirical community-level data from lake Erhai in China, where a critical transition was observed in 2001 [48]. Secondly, we calculate the neural network energy E_N on the same data for both test-cases in order to evaluate our theoretical expectations for E_N. For comparison with conventional leading indicators of critical transitions, we also calculate time-series metrics (TSMs: autoregressive coefficient at lag 1 (AR1), standard deviation (SD), skewness, and kurtosis), biodiversity, and for empirical data the sequential disorder-biodiversity correlation. Detailed methods and scripts are available in the Materials and Methods section. Both metrics find evidence of increasing diatom community instability during eutrophication in each dataset, with eco-net energy revealing complex eco-memory dynamics occurring before and during the regime shift.

Results and Discussion

Model data: PCLake+

We first apply the algorithm to output from a default setup of PCLake+ [49], an extension of the widely-used PCLake model of lake eutrophication, as a test-bed with well-known dynamics and drivers for generating realistic data. The artificial lake supports 14 functional groups of species, and has phosphorus input increasing along a nonlinear but non-hysteretic load-response curve to induce eutrophication (detailed in Materials and Methods). The impact of nutrient enrichment by phosphorus is clearly visible across three distinct phases in lake conditions, dominant eigenvalue (λ_d; i.e. local instability), eco-net energy (E_N), and biodiversity (Figure 4a-d, left). During Phase 1, initial destabilisation, λ_d and E_N both increase after the start of nutrient input and the replacement of vegetation by plankton, with both peaking sharply at the phase's end along with biodiversity. At the beginning of Phase 2, pre-transition, declining zooplankton abundance is marked by falls in both λ_d and E_N, followed by both plateauing with the continued rise in phytoplankton (indicated by chlorophyll concentration). Phase 3, regime shift, is marked by initial peaks followed by steady declines in both λ_d and E_N, with biodiversity continuing to decline towards eutrophic conditions.

These trends in λ_d and E_N, and the fluctuations around them, indicate complex eco-memory dynamics starting decades before the transition. The initial destabilisation away from memorised conditions, reflected in the rising λ_d indicative of weakened net-negative feedbacks, was followed by multi-phase "relearning" processes, reflected in decreasing E_N, as the eco-net reconfigured its dominant functional groups. Because PCLake+ was initialised and run as a stable non-eutrophic system prior to exogenous forcing by nutrient input, the system initially had only this one memorised attractor in its energy landscape. The observation of declining E_N after an initial destabilisation phase indicates learning of a new and more stable attractor (as conceptualised in Figure 3c). Increasing the sediment

accumulation rate through the artificial core does not alter the overall signal but reduces the temporal resolution of deeper signals (Supplementary Figure S1), indicating this methodology is not overly sensitive to temporal resolution.

Empirical data: lake Erhai

Similar to our PCLake+ scenario, lake Erhai in south-western China transitioned abruptly to sustained eutrophication after decades of exogenous forcing from phosphorus input, which were largely absorbed by lake sediment prior to the transition. Diatom abundances were sampled at regular radioisotope-dated intervals down sediment cores [28,48]. Because they are well preserved through time, the time-series of diatom composition provides a useful indicator of ecosystem state for producer-phytoplankton functional groups. We posit that the community-level λ_d for diatoms acts as a proxy for whole-ecosystem resilience, given that diatoms play a key role in the trophic loops involved in eutrophication [30] and are ecologically sensitive to water quality [50,51]. Despite not having data on the whole ecosystem network, in common with almost all empirical studies, we can still expect λ_d and E_N to record changes in wider network connections during a major regime shift that impacts on the whole ecosystem.

We observe comparable phases of activity to PCLake+ (Figure 4a-d, right). Phase 1, initial destabilisation, begins with the start of nutrient enrichment ~50-60 ya, as reported in historical records [48]. This phase is marked by an increasingly negative biodiversity-disorder correlation in diatom composition, indicative of strengthening interactions between species [27], accompanied by rising E_N towards a brief peak in λ_d. In Phase 2, pre-transition, λ_d sustains a low plateau and E_N drops to a low fluctuating average, while the biodiversity-disorder correlation continues its increasingly negative trend. Just prior to the observed transition at ~8 ya, E_N increases rapidly to its highest level. In Phase 3, regime shift, both E_N and λ_d drop slightly, accompanying the sharp drop in biodiversity and the return to positive

disorder-biodiversity correlation indicative of weakly interacting species [27].

These results show a similar pattern to the PCLake+ results. Phase 1 is marked by increasing E_N and a sharp peak in λ_d indicative of the eco-net destabilising decades before the transition. At the start of Phase 2, the sharp drop in E_N indicates relearning of newly forming eco-net configurations. Lower E_N could also represent the network shifting to a pre-existing alternative and deeper memory attractor, as lake Erhai may have experienced and memorised eutrophic conditions before (the reverse of Figure 3b). The subsequent rapid increase in E_N up to the transition into Phase 3 is consistent with a composition dominated by strongly interacting species for which the system has no previous memory. The post-transition decline in λ_d and E_N partly mirror PCLake+ but the decline is smaller for E_N, which may result from slower forcing (Supplementary Figure S2) or the full decline has yet to be observed. The overall similarity in phasing between lake Erhai and PCLake+ suggests that diatoms can act as a community-level proxy for the whole lake ecosystem. However, the progression of λ_d and E_N prior to Phase 1 also shows marked changes in magnitude, suggesting that these metrics respond also to events unrelated to the eventual regime shift.

Comparison to time-series metrics

TSMs are shown for comparison as a conventional resilience indicator method. As a consistently sampled model lake with constant temporal resolution, PCLake+ provides idealised conditions for observing TSMs during eutrophication (Figure 4e). Prior to a critical transition autocorrelation (AR1) and variability (SD) are expected to increase as a result of 'critical slowing down', while kurtosis may also increase and skewness may change either way [6]. Phase 1 has rapidly peaking AR1, skewness, and kurtosis, while SD begins a steady increase. Phase 2 has declines in all metrics, except for SD which continues its steady rise. Lake Erhai shows the largest changes in metric magnitudes prior to Phase 1 at ~90 ya, which may represent an unrecorded precursor event. Its Phase 1 nevertheless has rising AR1,

skewness and SD as seen in PCLake+, but a steady decline in kurtosis (in contrast to PCLake+). Its Phase 2 has declines in skewness and kurtosis while SD continues to increase (consistent with PCLake+), but AR1 stays high (in contrast to PCLake+). Both examples suggest SD consistently increases during destabilisation, but variable temporal resolution can obscure this in real-world data (Supplementary Figure S1). However, due to the methodological limitations described earlier, TSM are not considered robust for lake Erhai without further sensitivity and significance testing, and are shown for comparison only.

Further Development

Future iterations of these metrics for resilience and ecological memory should aim to address some current limitations. The network-inference algorithm functions most efficiently on long datasets that sufficiently resolve pre-perturbation equilibrium population dynamics and had sensitive detection thresholds to minimise data sparseness, but few palaeorecords meet this ideal. Recent developments in metagenomic network inference [38] suggest scope for improving efficiency on short or sparse data. We have assumed that ecosystem dynamics follow a generalised Lotka-Volterra model, in which all abundance changes are caused by linear pairwise interspecific interactions, and all other processes are pooled into a noise term. However, nonlinear interactions are expected in some ecosystems [37,52–56] and potentially allow a better representation of multiple alternative stable states [57,58], but are harder to parameterise. Future work will assess the feasibility of allowing nonlinear functional responses. The model currently includes only biotic interactions, with life-environment interactions implicit in parameter constants. Incorporating life-environment interactions explicitly would allow for more realistic feedback loops to emerge (such as anoxia-driven phosphorus release from sediment) that are critical for understanding resilience. However, the eco-net cannot simply be extended to include abiotic elements, because the environment is assumed to be its training input. Possible solutions include neural networks that allow training

input that is dynamic (e.g. continuous-time recurrent neural networks) and interacts with the eco-net itself (e.g. multi-layer or adversarial networks).

Conclusions

Our analyses indicate that changes in past ecosystem resilience and memory can be reconstructed in palaeoecological datasets of community composition from various settings, providing complementary information to metrics-based early warning signals. Our development here of eco-net energy has allowed us to explore eco-memory in process-based models and empirical data for the first time. Further work is required to fully understand the drivers and implications of eco-net energy dynamics, and to disentangle the effects of eco-memory from other drivers of ecosystem resilience. Eco-memory opens up a new dimension for understanding ecosystem resilience, with the formation of eco-memory potentially increasing resilience by allowing past stable eco-network states to be recovered after disruptions.

Materials and Methods

Ecological network inference

All analysis was conducted in R [47]. Abundance data – in the form of the relative proportional abundance per species at regularly sampled intervals from either empirical or model sources – were first transformed by isometric log ratio to account for compositionality in relative data and imputed to counter data sparseness [59] to form abundance matrix $\boldsymbol{X}$. To find the structure of the eco-net, we followed the methods of metagenomics network inference [39,40]. We first assumed that ecological interactions follow a generalised Lotka-Volterra model:

$$\frac{dx_i(t)}{dt} = x_i(t)\left[\beta_i - \sum_j \alpha_{ij} x_j(t)\right]$$

Equation 1

where x is the abundance of species i or j, α_{ij} is the interaction coefficient between species i and j ($<-1 < \alpha_{ij} < 1$), and β_i is the constant unitary net growth rate of species i. Taking the logarithm of a discrete-time Lotka-Volterra model based on Equation 1 and assuming constant time-steps [39] gives:

$$ln\, x_i\,(t+1) - ln\, x_i\,(t) = \zeta_i(t) + \beta_i - \sum_j \alpha_{ij}(x_j(t) - \langle x_j \rangle)$$

Equation 2

where ζ is the noise term and $\langle x_j \rangle$ is the median abundance (approximating to equilibrium abundance and carrying capacity for normalising against; assumes sufficient data to capture equilibrium population dynamics and ephemeral species). Given abundance time-series, the interaction coefficients α_{ij} can be found by performing multiple linear regression (lasso regularised due to data sparseness and limited observations relative to variables, and k-Fold cross-validated with at least 2 iterations to improve prediction performance) of $[ln\, x_i\,(t+1) - ln\, x_i\,(t)]$ against $\left[x_j(t) - \langle x_j \rangle\right]$. The regression coefficients then populate the elements of interaction matrix $\boldsymbol{A}$, which represents the eco-net's structure at time t. The interaction matrix was calculated from the abundance data on a 50% rolling window in order to capture changes in the eco-network over time. We use non-interpolated data to avoid well-known biases interpolation introduces to time-series analysis [26], although this results in slightly varying time-steps. Network inference and linear regression are less sensitive to non-equidistant time points than time-series analysis [38], but the discrete-time Lotka-Volterra model still assumes constant time-steps. To minimise this limitation, input data should be scrutinised prior to analysis and sections with substantial shifts in time resolution excluded.

Local Stability Analysis

The eco-net's stability was estimated using local stability analysis, by taking the absolute of the interaction matrix $\boldsymbol{A}$ (which is equivalent to the Jacobian of the Lotka-Volterra

system [30–32]; the system is considered stable at its equilibrium point if the real part of the Jacobian's eigenvalues remain negative, indicating net negative feedbacks), setting intraspecific interaction (α_{ii}) to 0 to satisfy the Perron-Frobenius Theorem [33], and calculating the dominant eigenvalue λ_d of the new positive matrix $\underline{\boldsymbol{A}}$. The Jacobian's λ_d is tightly related to other ecosystem stability metrics, having been shown to be equivalent to the smallest value for α_{ii} necessary for the interaction matrix $\boldsymbol{A}$ to be quasi-diagonally dominant and therefore stable, as well as being related to the maximum loop weight of the longest trophic loop [30,33–35]. λ_d was calculated on the rolling interaction matrix in order to estimate changes in ecosystem stability over time.

Ecological memory and eco-net energy

Recent theory suggests that ecosystems may have a distributed 'memory' of past states as a result of a process akin to unsupervised Hebbian learning in a Hopfield network [41,42] (see Introduction). There is no direct metric of Hopfield network memory strength, but we can calculate its energy, which is minimised at stable points. By assuming the eco-net's interaction matrix $\boldsymbol{A}$ is equivalent to a neural network's weight matrix $\boldsymbol{W}$ [41] and treating it as a continuous Hopfield network, we posit the eco-net energy at time t, E_N, can be calculated by:

$$\mathrm{E}_N = -\frac{1}{2}\mathbf{X}(\mathbf{t})^{\boldsymbol{T}}\mathbf{A}\mathbf{X}(\mathbf{t})$$

Equation 3

where $\boldsymbol{X}(\boldsymbol{t})$ is a vector of species abundances at time t [43–46]. We expect low E_N for eco-nets that have 'learnt' from stable environmental conditions or have shifted to a more stable attractor (Figure 3a), and higher E_N when destabilisation shifts the eco-net away from its learned state or to a less stable attractor (Figure 3b). However, asymmetric matrices with non-zero diagonals, such as an ecosystem interaction matrix, can lead to chaotic behaviour and the network converging to alternative metastable points, meaning eco-nets may not always

completely minimise their energy.

Comparison Metrics

For comparison we also calculated normalised time-series metrics (TSMs: Z-scores of AR1, standard deviation (SD), skewness, and kurtosis) on the abundance of a previously identified keystone species (*Cyclostephanos dubius*; Erhai [28,48]) or surface chlorophyll concentration (PCLake+) with variance-stabilising square-root transformation, Gaussian detrending (bandwidth: PCLake+ 1%; Erhai 5%), and a rolling window of 50% up to but not including the transition (observed critical transition for lake Erhai, nutrient-enrichment ceasing for PCLake+) (R: *generic_ews*, "*earlywarnings*"[6]). However, due to the methodological limitations described in the Introduction [7,23,24,26,27], time-series metrics are shown for comparison only. We also calculate compositional disorder metrics [28] (disorder/nestedness, biodiversity (inverse Simpson index), and their correlation using pre-transition window; R: *nestedtemp* & *diversity*, "*vegan*"[60]) where possible.

PCLake+ Modelling

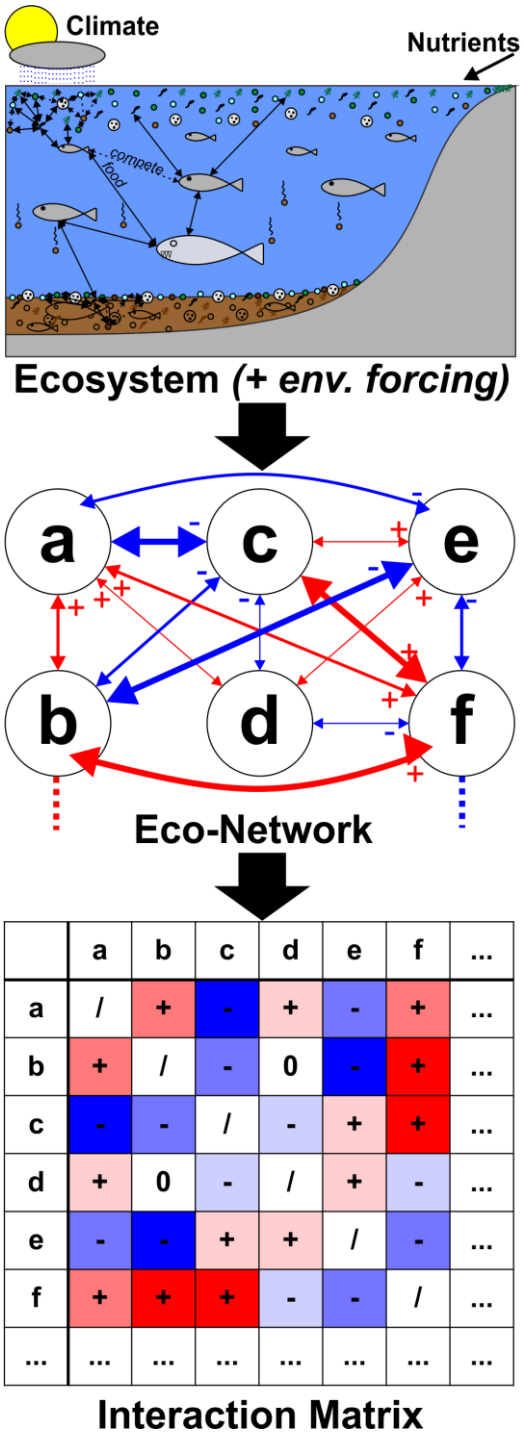
PCLake+ is initialised in its default clear deep state with monthly time-steps, run for 200 years to ensure stability, and nutrient input increased from 0.0001 to 0.005 $gPm^{-2}d^{-1}$ (with the latter value considered sufficient to induce a eutrophic regime shift in PCLake [61,62]) over 50 years to induce eutrophication. Eutrophication is visible in the results as a shift in dominance from vegetation to plankton, which is non-abrupt as PCLake+ in its default settings displays a nonlinear but non-hysteretic load-response curve [49,63]. Model outputs consist of dry-weight abundances for each functional group (diatoms, green algae, blue-green algae, vegetation, zooplankton, and detritus split into surface and benthic groups, along with fish split into adult, juvenile, and piscavorous). As real-world palaeodata has reduced resolution compared to the daily/monthly model output as a result of sedimentary

preservation, it was necessary to bin the model output into intervals equivalent to depth sampling in order to generate realistic artificial lake palaeorecords. This also allowed us to simulate the differing effects of constant or variable accumulation rates on our results (Supplementary Figure S1). The binned model outputs are then used as abundance inputs for the algorithm. We cannot compute nestedness or nestedness-biodiversity correlation for PCLake+ as only functional groups which always have some presence (rather than individual species which can go locally extinct) are available.

Acknowledgements

Erhai data was provided by the Nanjing Institute of Geography and Limnology. We thank Pete Langdon and Richard Watson for comments on preliminary results, Annette Janssen for advice on using PCLake+.

Figures



	a	b	c	d	e	f	...
a	/	+	-	+	-	+	...
b	+	/	-	0	-	+	...
c	-	-	/	-	+	+	...
d	+	0	-	/	+	-	...
e	-	-	+	+	/	-	...
f	+	+	+	-	-	/	...
...	...	...	...	...	...	...	...

Figure 1: Schematic illustrating how ecosystem interactions (top) are represented as an eco-network (middle –lettered nodes represent individual species or functional groups, and edges their positive (red) or negative (blue) interactions), the structure of which forms an interaction matrix (bottom) suitable for resilience analysis. Only symmetric interactions are shown in this schematic; the model can also accommodate asymmetric interactions such as asymmetries in competitive (-/-) and mutualistic (+/+) interactions, and predator-prey (+/-) and host-parasite (-/+) interactions.

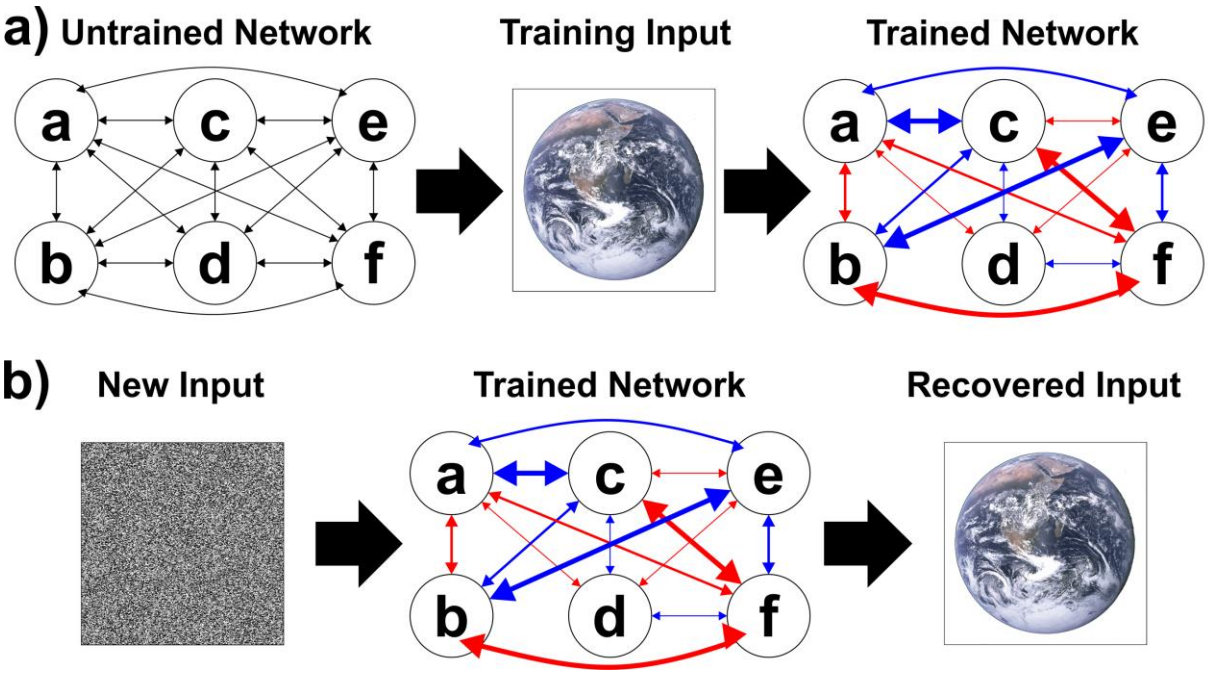


Figure 2: Schematic illustrating how (a) neural (or ecological) networks are trained by exogenous forcing (with environmental drivers), and (b) the character of this training input can be recovered after applying new degraded input to the trained network instead. In this case relearning is not enabled and so the new input does not lead to network alterations. Lettered nodes represent individual species or functional groups, and edges their positive (red) or negative (blue) interactions; black arrows in the untrained network indicate initial null or random values prior to training. Only symmetric interactions are shown in this schematic; the model can also accommodate asymmetric interactions such as asymmetries in competitive (-/-) and mutualistic (+/+) interactions, and predator-prey (+/-) and host-parasite (-/+) interactions.

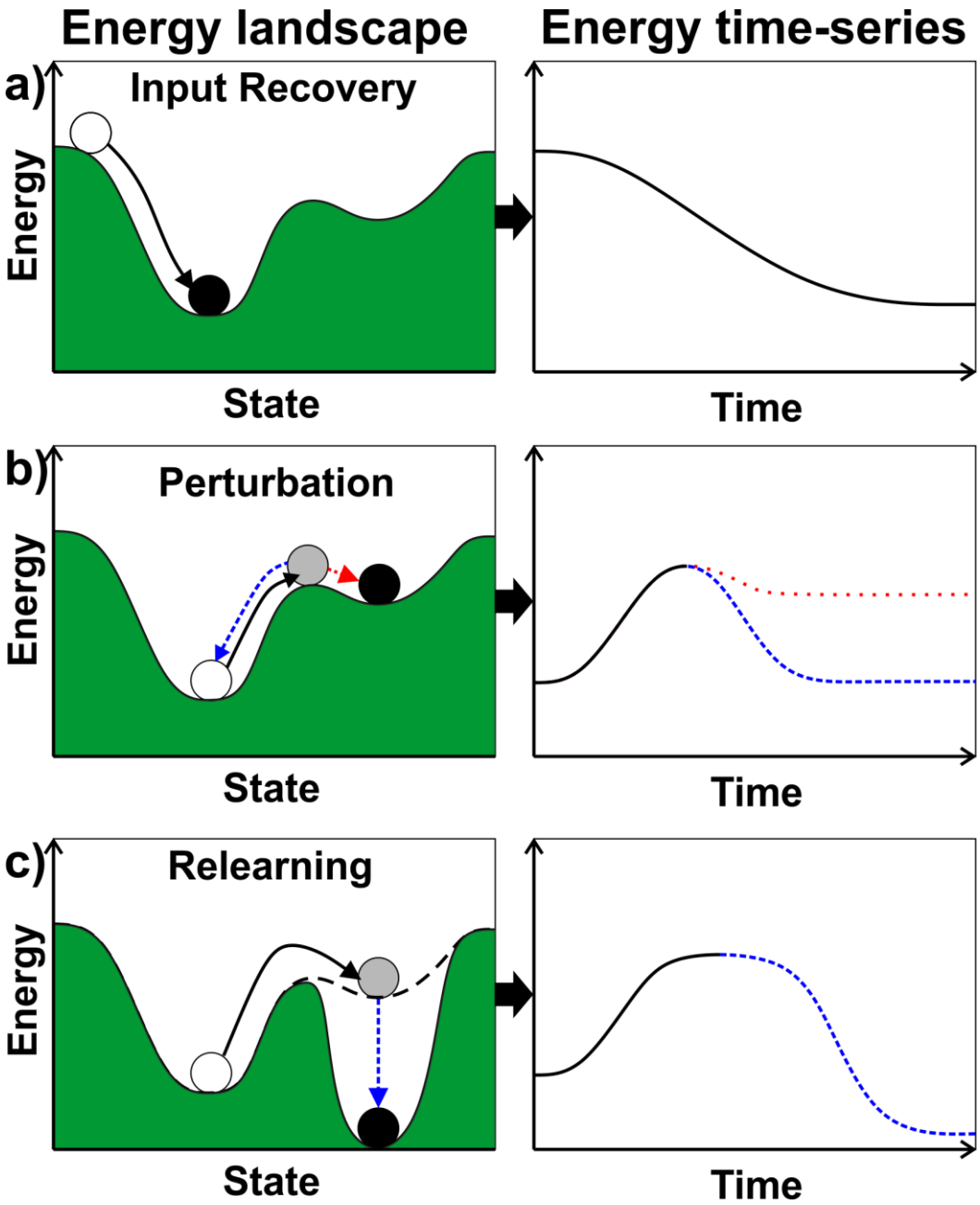


Figure 3: Schematic illustrating the dynamics of network energy in a neural network with attractor "memories". In (a) the neural network has formed an energy landscape (left) after learning from training input, allowing the stored memory to be recovered if the input is reset, as illustrated by the energy of the ball declining (right) as it moves into the energy minima (i.e. the memorised state). In (b), the state of the network is then perturbed (left) leading to a temporary increase in the network's energy (right), followed by the network either returning to the original attractor and low energy state (dashed blue arrow/line) or passing into an alternative, less stable attractor with higher energy (i.e. a different, weaker memory – dotted red arrow/line), In (c) network relearning is possible, and the energy landscape adjusts to allow the alternative attractor/memory to deepen in the network's new state (left, dashed blue arrow), leading to the network eventually reaching an even lower energy than before (right, dashed blue line).

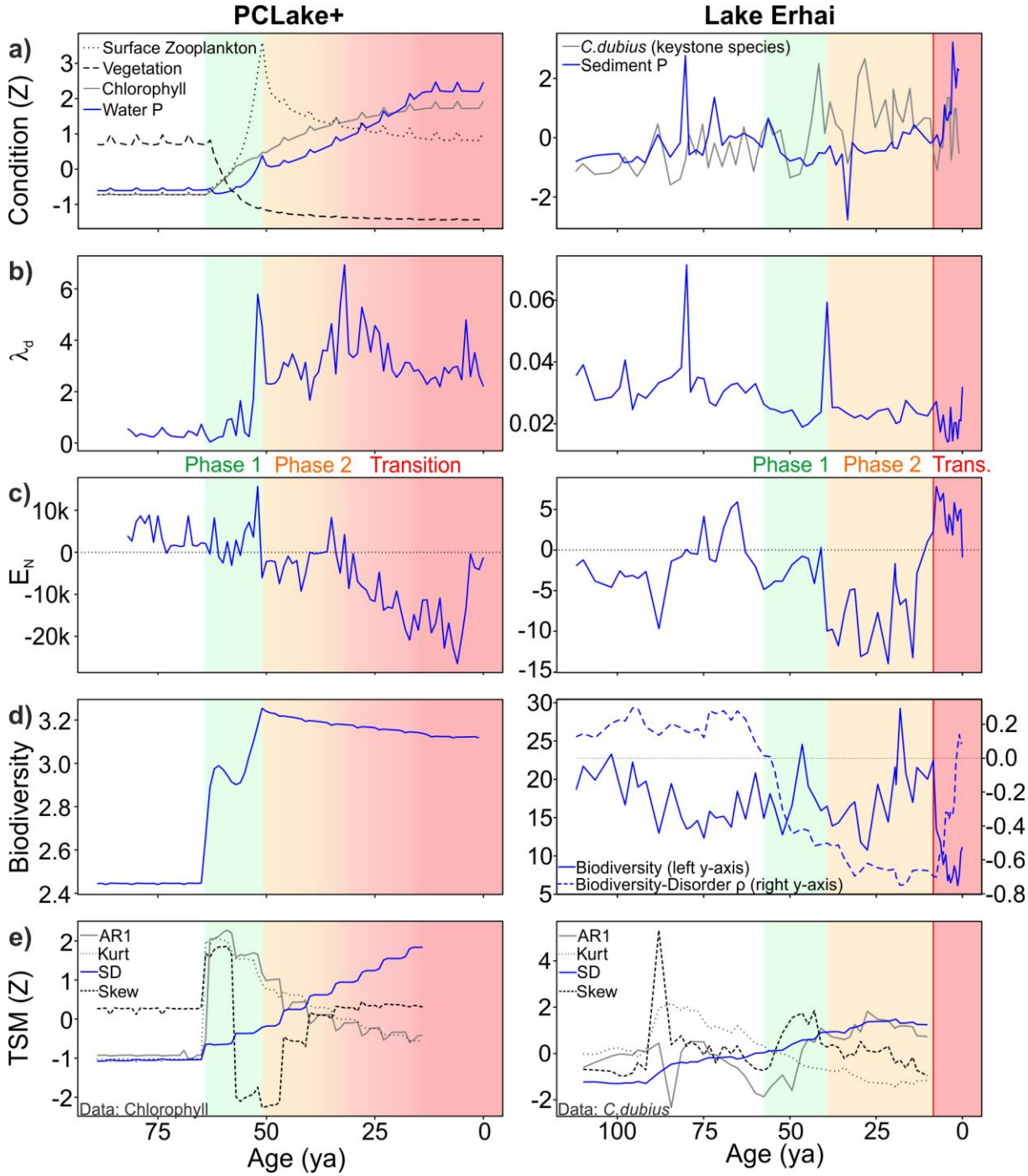


Figure 4: Time-series of lake conditions and resilience metrics for PCLake+ (left) and lake Erhai (right) through time (years ago) under increasing concentrations of phosphorus (P) leading to eutrophication. (a) Lake conditions indicated by normalised concentrations/abundances of vegetation, zooplankton, and chlorophyll (phytoplankton) in PCLake+, and increased keystone species (*Cyclostephanos dubius*) abundance in Erhai; (b) dominant eigenvalue (i.e. local instability) λ_d; (c) eco-net energy E_N; (d) biodiversity (inverse Simpson index), and disorder-biodiversity correlation (Erhai only, [28]); (e) normalised time-series metrics (TSM; for comparison only). Phase 1: initial destabilisation (green); Phase 2: pre-transition (orange); Phase 3: regime shift (red, red vertical line indicates known critical transition).

References

1. Scheffer M, Bascompte J, Brock WA, Brovkin V, Carpenter SR, Dakos V, et al. Early-warning signals for critical transitions. Nature. 2009;461: 53–9. doi:10.1038/nature08227

2. Dakos V, Carpenter SR, van Nes EH, Scheffer M. Resilience indicators: prospects and limitations for early warnings of regime shifts. Philos Trans R Soc B Biol Sci. 2015;370: 20130263–20130263. doi:10.1098/rstb.2013.0263

3. Lenton TM. Environmental Tipping Points. Annu Rev Environ Resour. 2013;38: 1–29. doi:10.1146/annurev-environ-102511-084654

4. Kéfi S, Dakos V, Scheffer M, Van Nes EH, Rietkerk M. Early warning signals also precede non-catastrophic transitions. Oikos. 2013;122: 641–648. doi:10.1111/j.1600-0706.2012.20838.x

5. Scheffer M, Carpenter SR, Lenton TM, Bascompte J, Brock WA, Dakos V, et al. Anticipating critical transitions. Science. 2012;338: 344–8. doi:10.1126/science.1225244

6. Dakos V, Carpenter SR, Brock WA, Ellison AM, Guttal V, Ives AR, et al. Methods for Detecting Early Warnings of Critical Transitions in Time Series Illustrated Using Simulated Ecological Data. Yener B, editor. PLoS One. 2012;7: e41010. doi:10.1371/journal.pone.0041010

7. Lenton TM. Early warning of climate tipping points. Nat Clim Chang. 2011;1: 201–209. doi:10.1038/nclimate1143

8. Dakos V, Scheffer M, van Nes EH, Brovkin V, Petoukhov V, Held H. Slowing down as an early warning signal for abrupt climate change. Proc Natl Acad Sci U S A. 2008;105: 14308–12. doi:10.1073/pnas.0802430105

9. Carpenter SR, Brock WA. Rising variance: a leading indicator of ecological transition. Ecol Lett. 2006;9: 311–318. doi:10.1111/j.1461-0248.2005.00877.x

10. Lenton TM, Livina VN, Dakos V, Scheffer M. Climate bifurcation during the last deglaciation? Clim Past. 2012;8: 1127–1139. doi:10.5194/cp-8-1127-2012

11. Scheffer M, Van Nes EH. Shallow lakes theory revisited: Various alternative regimes driven by climate, nutrients, depth and lake size. Hydrobiologia. 2007;584: 455–466. doi:10.1007/s10750-007-0616-7

12. Scheffer M, Hosper S, Meijer M, Moss B, Jeppesen E. Alternative equilibria in shallow lakes. Trends Ecol Evol. 1993;8: 275–279.

13. Scheffer M, Carpenter SR. Catastrophic regime shifts in ecosystems: Linking theory to observation. Trends Ecol Evol. 2003;18: 648–656. doi:10.1016/j.tree.2003.09.002

14. Carpenter SR, Ludwig D, Brock WA. Management of Eutrophication for Lakes Subject To Potentially Irreversible Change. Ecol Appl. 1999;9: 751–771. doi:10.1890/1051-0761(1999)009[0751:MOEFLS]2.0.CO;2

15. Carpenter SR. Eutrophication of aquatic ecosystems: Bistability and soil phosphorus. Proc Natl Acad Sci. 2005;102: 10002–10005. doi:10.1073/pnas.0503959102

16. Armstrong McKay DI, Lenton TM. Reduced carbon cycle resilience across the Palaeocene–Eocene Thermal Maximum. Clim Past. 2018;14: 1515–1527. doi:10.5194/cp-14-1515-2018

17. Holling CS. Engineering resilience versus ecological resilience. Eng within Ecol constraints. 1996; 31–43. Available: http://books.google.com/books?hl=en&lr=&id=QE4ydbFti7AC&oi=fnd&pg=PA31&dq=Engineering+Resilience+versus+Ecological+Resilience&ots=28zLqqlAvu&sig=WgX_0iLQ9jQ4EqFLbgqxBY2aBYE

18. Gsell AS, Scharfenberger U, Özkundakci D, Walters A, Hansson L-A, Janssen ABG, et al. Evaluating early-warning indicators of critical transitions in natural aquatic ecosystems. Proc Natl Acad Sci. 2016;113: E8089–E8095. doi:10.1073/pnas.1608242113

19. Bestelmeyer BT, Ellison AM, Fraser WR, Gorman KB, Holbrook SJ, Laney CM, et al. Analysis of abrupt transitions in ecological systems. Ecosphere. 2011;2: art129. doi:10.1890/ES11-00216.1

20. Burthe SJ, Henrys PA, Mackay EB, Spears BM, Campbell R, Carvalho L, et al. Do early warning indicators consistently predict nonlinear change in long-term ecological data? J Appl Ecol. 2016;53: 666–676. doi:10.1111/1365-2664.12519

21. Hastings A, Wysham DB. Regime shifts in ecological systems can occur with no warning. Ecol Lett. 2010;13: 464–472. doi:10.1111/j.1461-0248.2010.01439.x

22. Litzow MA, Hunsicker ME. Early warning signals, nonlinearity, and signs of hysteresis in real ecosystems. Ecosphere. 2016;7. doi:10.1002/ecs2.1614

23. Boettiger C, Ross N, Hastings A. Early warning signals: The charted and uncharted territories. Theor Ecol. 2013;6: 255–264. doi:10.1007/s12080-013-0192-6

24. Boettiger C, Hastings A. Early warning signals and the prosecutor's fallacy. Proc Biol Sci. 2012;279: 4734–9. doi:10.1098/rspb.2012.2085

25. Dakos V, Carpenter SR, van Nes EH, Scheffer M. Resilience indicators: prospects and limitations for early warnings of regime shifts. Philos Trans R Soc B Biol Sci. 2015;370: 20130263–20130263. doi:10.1098/rstb.2013.0263

26. Carstensen J, Telford RJ, Birks HJB. Diatom flickering prior to regime shift. Nature. 2013;498: E11-2. doi:10.1038/nature12272

27. Taranu ZE, Carpenter SR, Frossard V, Jenny J-P, Thomas Z, Vermaire JC, et al. Can we detect ecosystem critical transitions and signals of changing resilience from paleo-ecological records? Ecosphere. 2018;9: e02438. doi:10.1002/ecs2.2438

28. Doncaster CP, Chávez VA, Viguier C, Wang R, Zhang E, Dong X, et al. Early warning of critical transitions in biodiversity from compositional disorder. Ecology. 2016;97: 3079–3090. doi:10.1002/ecy.1558

29. Wang R, Dearing JA, Doncaster CP, Yang X, Zhang E, Langdon PG, et al. Network

parameters quantify loss of assemblage structure in human-impacted lake ecosystems. Glob Chang Biol. 2019; 1–12. doi:10.1111/gcb.14776

30. Kuiper JJ, van Altena C, de Ruiter PC, van Gerven LPA, Janse JH, Mooij WM. Food-web stability signals critical transitions in temperate shallow lakes. Nat Commun. 2015;6: 7727. doi:10.1038/ncomms8727

31. de Ruiter PC, Neutel A-M, Moore JC. Energetics, Patterns of Interaction Strengths, and Stability in Real Ecosystems. Science (80-). 1995;269: 1257–1260. doi:10.1126/science.269.5228.1257

32. Lade SJ, Gross T. Early Warning Signals for Critical Transitions: A Generalized Modeling Approach. PLoS Comput Biol. 2012;8: e1002360. doi:10.1371/journal.pcbi.1002360

33. Neutel A-M, Heesterbeek JAP, Ruiter PC de. Stability in Real Food Webs: Weak Links in Long Loops. Science (80-). 2002;296: 1120–1123. doi:10.1126/science.1068326

34. Neutel A-M, Heesterbeek JAP, van de Koppel J, Hoenderboom G, Vos A, Kaldeway C, et al. Reconciling complexity with stability in naturally assembling food webs. Nature. 2007;449: 599–602. doi:10.1038/nature06154

35. Francis TB, Wolkovich EM, Scheuerell MD, Katz SL, Holmes EE, Hampton SE. Shifting Regimes and Changing Interactions in the Lake Washington, U.S.A., Plankton Community from 1962–1994. Hazen EL, editor. PLoS One. 2014;9: e110363. doi:10.1371/journal.pone.0110363

36. Rawcliffe R, Sayer CD, Woodward G, Grey J, Davidson TA, Iwan Jones J. Back to the future: Using palaeolimnology to infer long-term changes in shallow lake food webs. Freshw Biol. 2010;55: 600–613. doi:10.1111/j.1365-2427.2009.02280.x

37. Faust K, Raes J. Microbial interactions: From networks to models. Nat Rev Microbiol. 2012;10: 538–550. doi:10.1038/nrmicro2832

38. Faust K, Lahti L, Gonze D, de Vos WM, Raes J. Metagenomics meets time series analysis: Unraveling microbial community dynamics. Curr Opin Microbiol. 2015;25: 56–66. doi:10.1016/j.mib.2015.04.004

39. Fisher CK, Mehta P. Identifying Keystone Species in the Human Gut Microbiome from Metagenomic Timeseries Using Sparse Linear Regression. PLoS One. 2014;9: 1–10. doi:10.1371/journal.pone.0102451

40. Mounier J, Monnet C, Vallaeys T, Arditi R, Sarthou A-S, Helias A, et al. Microbial Interactions within a Cheese Microbial Community. Appl Environ Microbiol. 2008;74: 172–181. doi:10.1128/AEM.01338-07

41. Power DA, Watson RA, Szathmary E, Mills R, Powers ST, Doncaster CP, et al. What can ecosystems learn? Expanding evolutionary ecology with learning theory. Biol Direct. 2015;10: 69. doi:10.1186/s13062-015-0094-1

42. Watson RA, Szathmáry E. How Can Evolution Learn? Trends Ecol Evol. 2016;31: 147–157. doi:10.1016/j.tree.2015.11.009

43. Hopfield JJ. Neural networks and physical systems with emergent collective

computational abilities. Proc Natl Acad Sci. 1982;79: 2554–2558. doi:10.1073/pnas.79.8.2554

44. MacKay DJC. Information Theory, Inference, and Learning Algorithms. Cambridge University Press; 2005. doi:10.1198/jasa.2005.s54

45. Rojas R. Neural Networks: a Systematic Introduction. 1996. doi:10.1016/0893-6080(94)90051-5

46. Kröse B, Smagt P van der. An Introduction To Neural Networks. 1996.

47. R Foundation for Statistical Computing. R (v.3.5.1): A language and environment for statistical computing. Vienna, Austria; 2016. Available: https://www.r-project.org/

48. Wang R, Dearing JA, Langdon PG, Zhang E, Yang X, Dakos V, et al. Flickering gives early warning signals of a critical transition to a eutrophic lake state. Nature. 2012;492: 419–22. doi:10.1038/nature11655

49. Janssen ABG, Teurlincx S, Beusen AHW, Huijbregts MAJ, Rost J, Schipper AM, et al. PCLake+: A process-based ecological model to assess the trophic state of stratified and non-stratified freshwater lakes worldwide. Ecol Modell. 2019;396: 23–32. doi:10.1016/j.ecolmodel.2019.01.006

50. Hall R, Smol J. Diatoms as indicators of lake eutrophication. In: Stoermer EF, Smol JP, editors. The Diatoms: Applications for the Environmental and Earth Sciences. Cambridge, UK: Cambridge University Press; 1999. pp. 128–168.

51. Dixit SS, Smol JP, Kingston JC, Charles DF. Diatoms: Powerful Indicators of Environmental Change. Environ Sci Technol. 1992;26: 22–33. doi:10.1021/es00025a002

52. Neutel A-M, Thorne MAS. Linking saturation, stability and sustainability in food webs with observed equilibrium structure. Theor Ecol. 2016;9: 73–81. doi:10.1007/s12080-015-0270-z

53. Suzuki K, Yoshida K, Nakanishi Y, Fukuda S. An equation-free method reveals the ecological interaction networks within complex microbial ecosystems. McMahon S, editor. Methods Ecol Evol. 2017;8: 1774–1785. doi:10.1111/2041-210X.12814

54. Xiao Y, Angulo MT, Friedman J, Waldor MK, Weiss ST, Liu Y-Y. Mapping the ecological networks of microbial communities. Nat Commun. 2017;8: 2042. doi:10.1038/s41467-017-02090-2

55. Levine JM, Bascompte J, Adler PB, Allesina S. Beyond pairwise mechanisms of species coexistence in complex communities. Nature. 2017;546: 56–64. doi:10.1038/nature22898

56. Cao H-T, Gibson TE, Bashan A, Liu Y-Y. Inferring human microbial dynamics from temporal metagenomics data: Pitfalls and lessons. BioEssays. 2016;39: 1600188. doi:10.1002/bies.201600188

57. Gunderson LH. Ecological Resilience—In Theory and Application. Annu Rev Ecol Syst. 2000;31: 425–439. doi:10.1146/annurev.ecolsys.31.1.425

58. Kuiper JJ, Kooi BW, Peterson GD, Mooij WM. Bridging theories for ecosystem stability through structural sensitivity analysis of ecological models in equilibrium. bioRxiv. 2019; 1–23. doi:10.1101/2019.12.24.887901

59. Aitchison J. The Statistical Analysis of Compositional Data. J R Stat Soc Ser B Methodol. 1982;44: 139–177. doi:10.2307/2345821

60. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, et al. vegan: Community Ecology Package. 2017. Available: https://cran.r-project.org/package=vegan

61. Janse JH, De Senerpont Domis LN, Scheffer M, Lijklema L, Van Liere L, Klinge M, et al. Critical phosphorus loading of different types of shallow lakes and the consequences for management estimated with the ecosystem model PCLake. Limnologica. 2008;38: 203–219. doi:10.1016/j.limno.2008.06.001

62. NIOO. Manual PCLake, Version 6.13.16.2: DATM implementation in Microsoft Excel. 2017.

63. Janssen ABG, de Jager VCL, Janse JH, Kong X, Liu S, Ye Q, et al. Spatial identification of critical nutrient loads of large shallow lakes: Implications for Lake Taihu (China). Water Res. 2017;119: 276–287. doi:10.1016/j.watres.2017.04.045