Ecological Complexity 47 (2021) 100944

Contents lists available at ScienceDirect

Ecological Complexity

journal homepage: www.elsevier.com/locate/ecocom

Early warning signals of regime shifts for aquatic systems: Can experiments help to bridge the gap between theory and real-world application?

Julio Alberto Alegre Stelzer [a,b,c,*], Jorrit Padric Mesman [a,d], Rita Adrian [b,c], Bastiaan Willem Ibelings [a]

[a] *Department F.-A Forel for Environmental and Aquatic Sciences & Institute for Environmental Sciences, University of Geneva, 66 Blvd. Carl Vogt, 1205 Geneva, Switzerland*

[b] *Department of ecosystem research, Leibniz-Institute of Freshwater Ecology and Inland Fisheries, Muggelseedam, 301 Berlin, Germany*

[c] *Department of Biology, Chemistry, and Pharmacy, Freie Universität Berlin, Königin-Luise-Straße 28-30, Berlin, Germany*

[d] *Department of Ecology and Genetics, Uppsala University, Norbyvägen 18D, 75236 Uppsala, Sweden*

A R T I C L E I N F O

Keywords:
Early warning signals of a regime shift
Ecosystem resilience
Network ecology
Ecological forecast
Regime shifts
Preventive management

A B S T R A C T

Early-warning signals of a regime shift (EWS) indicate, for a wide range of systems, if a tipping-point is being approached. In ecology, EWS are well established from a theoretical perspective but are far from unequivocal when applied to field data. The gap between theory and application is caused by a set of limitations, like the lack of coherence between different EWS, data acquisition issues, and false results. Experiments assessing EWS may provide an empirical mechanistic understanding of why an EWS was observed (or failed to be observed), which often cannot be elucidated by simple computational modeling or pure environmental data. Here we focused on aquatic experiments to explore to what extent the existing EWS experiments can bridge the gap between the theory and real-world application. For that, we used the Thomson-ISI Web of Science© database to retrieve EWS experiments executed before early-2020, detailing their experimental designs and each EWS assessed. Success rates - correct anticipation of tipping points – were around 70% for the most used EWS (assessment of autocorrelation, variance, recovery, and shape of the distribution using abundance, Chlorophyll-a, Phycocyanin, and dissolved oxygen data). Yet, no EWS showed to be 100% reliable, and their use demands cautious interpretation. As a rule, we observed that experiments were not designed to tackle issues encountered in real-world situations. They lack a deep mechanistic understanding of why, when, and how an EWS was observed or not. When experiments did aim to assess issues encountered in the real world, the experimental designs were often of low ecological significance. We also investigated the relationship between sampling and the success rate of EWS, observing that the sampling regime might have to be tailor-made towards specific monitoring objectives. Moreover, experiments have taught us that the use of EWS can be more versatile than expected, going from monitoring the extinction of single populations to the anticipation of transient regime shifts. Most of the experiments presented here supported empirical proof of the existence of EWS in aquatic systems. Still, to bridge the gap between theory and application, experiments will have to move closer to real-world conditions and better support a mechanistic understanding of why EWS may succeed or fail to anticipate a regime shift. For that, we provide six elements to take into account when designing experiments that could enhance the capabilities of EWS to go beyond the stage of proof-of-concept.

1. Introduction

Studies from different scientific fields have suggested the existence of generic early-warning signals of a regime shift (EWS) that may indicate if a tipping-point is being approached (Scheffer et al., 2009). EWS are statistical signatures of system dynamics that change with proximity to a critical transition. For instance, a given response variable is predicted to increase in autocorrelation and variance but decrease its recovery rate after perturbations when the system approaches a critical transition (Dakos et al., 2012). These changes are often, but not always, a

* Corresponding author.
E-mail addresses: Julio.AlegreStelzer@unige.ch, jaastelzer@gmail.com (J.A.A. Stelzer), Jorrit.Mesman@unige.ch (J.P. Mesman), adrian@igb-berlin.de (R. Adrian), Bastiaan.Ibelings@unige.ch (B.W. Ibelings).

https://doi.org/10.1016/j.ecocom.2021.100944
Received 25 January 2021; Received in revised form 4 June 2021; Accepted 11 June 2021
Available online 9 July 2021

consequence of a phenomenon called *critical slowing down* that has been associated to loss of system's resilience (Scheffer et al., 2009).

In ecology, efforts have been made to identify proximity to tipping points, and after over a decade of development, the use of EWS are well consolidated from a theoretical and model perspective (e.g., Brock and Carpenter (2010); Carpenter et al. (2008); Dakos et al. (2012); Scheffer et al. (2012); van Nes and Scheffer (2007)). However, when assessing EWS in field data (the "real world") and questioning their potential use for the management and conservation of ecosystems, the interpretation of EWS are far from unequivocal. Some studies have reported that EWS failed in anticipating well-stated regime shifts (e.g., Bestelmeyer et al. (2011)); others found a mixture of successes and failures (e.g., Burthe et al. (2016); Gsell et al. (2016); Krkosek and Drake (2014); Litzow et al. (2013)); while yet others presented evidence supporting the theory of EWS (e.g., Wouters et al. (2015)). These conflicting results are unsurprising when predictions derived from simple models are upscaled to complex ecosystems (Scheffer et al., 2015). However, this leaves the question open on how to apply EWS in real-world situations.

EWS are most likely to be observed when changes in a driving force slowly push the system towards a tipping point (Dakos et al., 2015), yet exceptions exist (Dakos et al. (2015); Kéfi et al. (2013)). In fact, EWS have a series of caveats that have to be addressed before their application. For instance, simulation studies have pointed out that EWS may fail to anticipate regime shifts due to (i) a low signal-to-noise ratio (Clements et al. (2015); Dakos et al. (2015); Perretti and Munch (2012)), (ii) a lack of sensitivity of the response variable (e.g., Dakos et al. (2011)), or (iii) the existence of false positives and negatives (e.g., Boettiger and Hastings (2012); Boettiger et al. (2013)). For field data, the limitations increase, and EWS are reported to be also strongly affected by (iv) our capability to select measurable variables that are relevant for whole-system analysis (Burthe et al., 2016), (v) a lack of coherence between different EWS indicators - (e.g., when different metrics applied with the same response variable present divergent results - see Gsell et al. (2016), Burthe et al. (2016)), (vi) the requirement for ecosystem-specific knowledge of transition-generating mechanisms (Gsell et al. (2016), Spears et al. (2017)), and (vii) data acquisition issues (e.g., collecting a timeseries of sufficient length and quality). In short, the theory of EWS and its application to field data poses a set of limitations that have to be overcome in order to apply them successfully to real-world situations. This is where experiments might be helpful.

Experiments can help make sense of real-world situations in which a high-level complexity is often prohibitive for a complete understanding. Simultaneously, experiments may provide an empirical mechanistic understanding of why a given EWS was observed (or failed to be observed), which often cannot be elucidated from simple computational modeling. In this way, experiments may help to indicate (i) which are the most sensible response variables, (ii) the most reliable statistical signatures, (iii) how to identify false positive and negatives within a system (presence of EWS but no regime shift, i.e., due to the lack of relevance of the response variable or low signal-to-noise ratio) and, (iv) what is the minimum data requirement for applying EWS methods.

Here we explore to what extent existing experiments can help bridge the gap between the theory and real-world application of EWS. For that, we focus on experiments in aquatic systems, both freshwater and marine. Over the last years, there has been an increase in the number of aquatic EWS experiments that could potentially link theory and application, including whole-ecosystem manipulations. We analyzed the literature to explore the following questions:

1) Experimental design:
   i) Were the experimental designs structured in a way to facilitate extrapolations to real-world scenarios?
   ii) Did the experiments rely on realistic driving forces and regime shifts that are also likely to occur in real-world situations?
   iii) Were the systems gradually pushed towards the tipping-point, ultimately forcing a regime shift, with or without additional disturbance?
2) Performance:
   i) Which were the most common categories of response variables used for observing EWS?
   ii) Which were the metrics of EWS with the highest success rates in experiments?
3) Methodology:
   i) Can actual experiments provide information on the ideal sampling regime for EWS?

Up to now, EWS have never been compiled into a comprehensive review that explores how experiments can contribute to bridging the gap between theory and application of EWS in aquatic ecosystems. We conclude with suggestions on how to design future experiments that could further support the application of EWS in real-world systems.

2. Methodology

2.1. Review of EWS in aquatic experiments

We used the Thomson-ISI Web of Science© database to select papers reporting the use of EWS in experiments starting from early-2020 and earlier. No clear link between environmental studies and EWS of regime shifts was found before the year 2000. We searched papers using the Boolean search query "TS= ("early warning*" OR "slowing down" OR "return time*" OR "return rate*" OR "recovery time" OR "recovery rate*" OR "engineering resilience") AND (experiment* OR environment* OR field) AND ("tipping point*" OR "regime shift*" OR "critical threshold*" OR bifurcation")". This generic search term returned 241 papers of which we screened their abstracts and selected papers dealing specifically with the experimental assessment of EWS in aquatic ecosystems (freshwater and marine). Model-based papers and purely observational data were not included in the review.

2.2. Categorization of the experiments and EWS success rate

First, we divided experiments according to their biological complexity. Assemblages composed of 1-2 species were called "simple" biological setups, while assemblages consisting of 3 or more species were called "complex" (simple and complex experiments). In this way, we separate phytoplankton cultures and simple predator-prey systems from more complex systems like cosms and whole-ecosystem manipulations. Next, for each experiment we extracted (i) which driving forces were used, (ii) how they were applied, (iii) what regime shift the system went through, (iv) which perturbations were used, and (v) how they were applied to the systems. For the used definitions of driving force, regime shift, and perturbation, see Box – Terminology.

Assessing an early-warning signal requires a timeseries of a quantitative response variable and a statistical analysis. The response variable (from here on called "proxy") was extracted from the original paper based on the nature of the data rather than the individual taxonomic group. That is, when data from different taxonomic groups (e.g., *Daphnia magna,* or grazers as a functional group) had the same nature (i.e., concentration of organisms), they were clustered into the same group (here, "abundance"). Clustering was needed due to the high taxonomic heterogeneity of the proxies. For each proxy, data was assessed using one or more statistical analyses (from here on called "metrics"). Like mentioned before, metrics were also grouped based on their nature (e.g., coefficient of variation, standard deviation, and variance were grouped into "variance"; skewness and kurtosis into "shape of distribution"; recovery time, recovery rate, and recovery path into "recovery"). Each combination of a proxy and a metric corresponds to the assessment of a unique EWS. When the authors described the EWS as capable of indicating proximity to the regime shift, we labeled them as "positive". When

EWS failed to anticipate the regime shift or the result was inconclusive, we labeled them as "negative". Further, we calculated success rates for each pairwise combination to assess how reliable each EWS was (number of positive results divided by the total number of assessments).

2.3. *Effect of sampling on success rates of EWS*

To assess the effect of sampling (number of sampling points and frequency of sampling) on success rates of EWS of simple and complex experiments, we fitted logistic regression models. We started from the most complex model, including an interaction between the number of sampling points and frequency of sampling. Model selection was based on Akaike Information Criterion (AIC) values. When AIC values differed more than two units between models, we selected the model with the lowest AIC. If the AIC were similar, we chose the simplest model that comprised at least two continuous variables (length, frequency, or sampling points). We used only the most frequently used combinations of proxy and metric since most of the combinations lacked sufficient data to test that they can be used as EWS (see session (2) Most reliable combination of proxies and metrics (EWS)). Pseudo-R^2 and p-value were calculated for the final models.

The metric "recovery" was assessed in a separate logistic regression model from the other metrics (autocorrelation, variance, the shape of the distribution). Recovery is calculated using only part of the data (from perturbation until recovery), and so it may have different data requirements that could muffle the results for other metrics.

All statistical analyses were done in R software V3.6.1 (R Core Team, 2020) using the package "MASS" (Ripley et al., 2013).

Box – Terminology

Perturbation – a temporary stress on the system that alters the arrangement of biotic and abiotic elements, pushing the system out of equilibrium. Some examples of perturbations are nutrient and flood pulses caused by a storm, heatwaves, or sudden mortality events (e.g., due to contamination, trolling). If the perturbation is too strong, a regime shift may unfold as a consequence of it.

Driving force – an inherent component of the system that pushes the system away from its actual ecological state. Common examples of changes in driving forces are oligo- or eutrophication, climate change-driven modifications in the physicochemical characteristics of the waterbody (temperature, pH, or physical structure of the water column), or changes in the relative importance of bottom-up and top-down control (abundance of top-predators). If the changes in the driving force are too sudden and with steep increase rates, it may also bring the system out of equilibrium, triggering a perturbation (e.g., sudden nutrient release from resuspended sediment).

Regime shift – a usually abrupt and often persistent change in the functioning of the ecosystem (ecological state). It can be caused either by changes in the driving force, due to a strong or additive perturbations, or the combination of both. Examples of well-known regime shifts are the transition from clear to turbid state in shallow lakes and from lentic to lotic states on rivers, but it can also represent the extinction of a species (going from present to absent). Regime shifts can be stable or transient depending on the strength of the existing self-sustaining feedbacks.

Ecological resilience – Capability of a system to sustain an ecological state when pushed either by a driving force or a strong perturbation (Holling, 1973). Sometimes the term is exchangeable with ecological stability or "resistance to changes in the state". Contrary to engineering resilience (see Pimm (1984)), it is rather more associated with a qualitative characteristic of the system (e.g., the shape of the basin of attraction as the capability to sustain a clear state of a lake under eutrophication).

For a more extensive in-depth view of the glossary proposed here, see Hodgson et al. (2015).

3. Results and Discussion

3.1. *Experimental Designs*

We found 24 scientific papers describing 19 unique aquatic experiments on EWS (Supplement material – List of Papers). Ten experiments were simple biological setups composed of 1-2 species, which renders them less useful to bridge the gap between theory and real-world application. Nevertheless, they are meaningful for providing consistent proofs-of-concept when transitioning from simple to more complex biological setups. The other nine experiments were complex biological setups and had the potential to mimic real-world situations. They include three whole-lake manipulations, one whole-river manipulation, four coastal manipulations, and one microbial ecosystem with five identified trophic levels. Below we scrutinize (a) the structure of these experimental designs and (b) the nature of the forces that induced regime shifts, pointing out the lessons learned for future use of EWS. The list of all regime shifts studies with ID numbers, and the number of EWS tested in each case is shown in Table 1.

3.1.1. *Structure of the experimental designs – increasing driving force vs. perturbations*

Two opportunities for assessing EWS are well described in the literature. First, when a continuous but slow increase in the driving force pushes the system towards a tipping point (Dakos et al., 2015). Second, when the above-mentioned scenario suffers multiple pulse perturbations, and the decrease in recovery is used to indicate loss of ecological resilience (van Nes and Scheffer, 2007). To see if experiments complied with these theoretical frameworks, we provide details on the driving forces and perturbations for each experiment (see Box – Terminology for the distinction between driving force and perturbation).

3.1.1.1. Simple biological setups. The simple biological setups tended to have driving forces that continuously increased over time, gradually pushing the systems towards the tipping point (six out of ten experiments - ID 2, 4, 10, 11, 13, 14). The other four simple experiments used categorical treatments to define different driving forces (e.g., temperature treatments (T) were T1= +2°C, T2= +3°C, T3= +4°C instead of a treatment starting on +0°C and ending at +4°C - (ID 5, 6, 12, 19). From a practical perspective, this scenario results in no real development towards a regime shift over time. The observation of regime shifts in these studies was done by comparing timeseries under different treatments rather than a continuous temporal assessment within the treatment itself. Such a design is a common and powerful experimental approach, but it has complications when upscaling to real environments. Selecting comparable natural systems that differ mainly in the level of a driving force (what would be analog to different treatments) is a real challenge - but a valuable one if we would like to rank systems in terms of resilience or examine systems in a gradient.

In nine of the ten simple experimental setups (ID 4, 5, 6, 10, 11, 12, 13, 14, 19), perturbations were applied. Perturbations created the opportunity for inferring a loss in ecological resilience by quantifying the recovery back to equilibrium. Recovery can be considered a metric based on the system's short-term responses compared to the classic long-term trends in metrics like autocorrelation and variance. Thus, it may be especially valuable when long-term monitoring data is not available. However, only five experiments combined perturbations with an increasing level in the driving force – the two most well-established conditions for assessing EWS (ID 4, 10, 11, 13, 14). In these five experiments, the driving force pushed the systems in the direction of a regime shift, whereas the perturbations moved the system further away from equilibrium. This combination of increasing driving force with independent perturbations is not paramount for identifying most EWS but consists of a scenario that comes closer to real-world situations. Ecosystems are continuously pressed by a multitude of driving forces (multiple stressors) while suffering stochastic perturbations that momentarily force them out of equilibrium (e.g., storms, heatwaves, point-source pollution events). Thus, this combination could be considered as a fruitful conceptual framework to be applied to experiments with near-natural scales.

3.1.1.2. Complex biological setups. With regards to the driving force, experiments with complex biological setups were the opposite of the simple ones. They were mainly based on setting treatments with steady-state driving forces instead of changing them over time (six out of nine -

Table 1
Summary of experiments used for observing EWS of regime shift in aquatic systems. Data reviewed in this paper comprise publications until early 2019. We total 119 EWS observations from 19 unique experiments, including from highly controlled experiments to whole-system manipulations. For individual EWS results, see Supplementary Material – Metadata. ID=Unique experimental ID, one experiment was described in more than one paper. Biological complexity: C – Complex; S – Simple biological setups. % Success = no.EWS described as capable of anticipating a regime shift in the original paper (positive result) divided by the total no.EWS.

ID	Perturbation	Regime of Perturbation	Driving Force	Is the Driving Force Increasing over Time?	no. EWS	% Success	Length of the Experiment	Regime Shift	**Biological Complexity**	**Reference**
1	Floods	Multiple Pulses	Time/n.a	No	9	67%	8 years	Flood Controlled / Flood Resilient Fauna	C - (Coastal Manipulation)	Robinson and Uehlinger (2008)
2	n.a	n.a	Food provision	Yes	4	100%	416 days	Extinction (Population Collapse)	S - (Microcosms)	Drake and Griffen (2010)
3	n.a*	Multiple Pulses	Top Predator Addition	Yes	28	57%	4 years	Planktivorous / Piscivorous	C - (Whole-Lake Manipulation)	Carpenter et al., (2011), Seekell et al. (2012), Batt et al. (2013), Pace et al. (2013), Seekell et al. (2013), Cline et al. (2014)
4	Mortality	Multiple Pulses	Light Intensity	Yes	3	67%	29 days	Extinction (Population Collapse)	S - (Single species setup)	Veraart et al. (2011)
5	Mortality[(a)]	Multiple Pulses	Mortality Rate	No	5	80%	9 days	Extinction (population collapse)	S - (Single species setup)	Dai et al., (2012)
6	Mortality	Multiple Pulses	Mortality Rate	No	4	75%	5 days	Extinction (population collapse)	S - (Single species setup)	Dai et al., (2013)
7	n.a*	Multiple Pulses	Organic matter	Yes	5	20%	4 days	Aerobic / Anaerobic	C - (Microcosms)	Sirota et al. (2013)
8	Mortality	Single Pulse	Nutrient Addition	No	1	0%	46 days	Extinction (Population Collapse)	C - (Coastal Manipulation)	Soissons et al. (2014)
9	Mortality	Single Pulse	Clipping	No	4	100%	7 years	Canopy/Turf dominated state	C - (Coastal Manipulation)	Benedetti-Cecchi et al. (2015)
10	Mortality	Multiple Pulses	Food provision	Yes	2	100%	22 days	Extinction (Population Collapse)	S - (Single species setup)	Dai et al. (2015)
11	Mortality	Multiple Pulses	Mortality Rate	Yes	2	100%	20 days	Extinction (Population Collapse)	S - (Single species setup)	Dai et al. (2015)
12	Mortality	Multiple Pulses	Temperature	No	3	100%	64 days	Extinction (Population Collapse)	S - (Single species setup)	Jarvis et al., (2016)
13	Mortality	Multiple Pulses	Food Provision	Yes	5	Indexed results	45 days	Extinction (Population Collapse)	S - (Microcosms)	Clements and Ozgul, (2016a)
14	Mortality	Multiple Pulses	Toxicity	Yes	6	33%	64 days	Extinction (population collapse)	S - (Microcosms)	Sommer et al. (2016)
15	n.a*	Single Pulse	Nutrient Addition	No	4	50%	2 years	Algal bloom	C - (Whole-Lake Manipulation)	Butitta et al., (2017)
16	Mortality	Multiple Pulses	Canopy Removal	No	6	67%	2 years	Canopy/Turf dominated state	C - (Coastal Manipulation)	Rindi et al., (2017), Rindi et al., (2018)
17	n.a	n.a	Nutrient Addition	Yes	26	62%	3 years	Algal bloom	C - (Whole-Lake Manipulation)	Wilkinson et al. (2018)
18	Mortality	Single Pulse	Tide Level	No	1	100%	2 years	Extinction (Population Collapse)	C - (Coastal Manipulation)	El-Hacen et al. (2018)
19	Mortality	Multiple Pulses	Mortality Rate	No	1	100%	16 days	Extinction (Population Collapse)	S - (Single species setup)	Ghadami et al. (2018)

*Perturbation was caused by the regime on which the driving force was changed [(a)]Secondary perturbation applied, Osmotic shock.

ID 1, 8, 9, 15, 16, 18). Only in three experiments the driving force increased over time (ID 3, 7, 17). These findings reflect the laborious and technically challenging process of modulating a specific driving force over an extended time in large-scale experiments. When using experiments to bridge the gap between theory and real-world application of EWS, the use of steady-state driving forces significantly differs from real-world scenarios. Under natural conditions, the driving forces are expected to change over time – pushing the system to the boundaries of its state - but most of the experiments using complex biological setups fail to incorporate this.

On the other hand, complex biological setups are often long-term experiments, taking months or even years to be accomplished. During this period, even when using a steady-state driving force, these experiments – when carried outdoors - are prone to environmental fluctuations. These fluctuations may encompass many different ecosystem stressors (e.g., water table variation – ID 1; nutrient oscillations – ID 3) or simply be caused by modifications on the steady-state condition established at the beginning of the experiment (e.g., due to nutrient leaching, interannual canopy growth). In fact, it is rare that the final conditions of a long-term experiment are comparable to its starting point. Thus, the length of the experiment may act itself as a composite of uncontrolled driving forces that changes the system over time (e.g., due to mineralization rates, biofilm formation, sludge formation). Nevertheless, when the driving force is not explicitly modulated, the experiment loses some of its potential for a mechanistic understanding of what triggered the ecological processes responsible for the regime shift. This is especially relevant in open systems where the driving force's maintenance or manipulation was done at infrequent intervals (e.g., annual clipping of canopy – ID 9; cover or seasonal addition of top predators – ID 3).

Concerning the perturbations, most of the complex biological setups were similar to the simple ones, imposing independent perturbations to the system (5 out of 9 - ID 1, 8, 9, 16, 18). In other cases, the perturbation was caused by an abrupt increase in the driving force that brought the system out of equilibrium (ID 3, 7, 15, see Box – Terminology, Perturbation, and Driving force). In those cases, recovery from perturbation could be quantified even without any specific additional manipulation. However, it is essential to notice that the recovery from the perturbation is entangled with the increase in the driving force in these cases. For instance, take an oligotrophic lake that receives a strong nutrient pulse. One may expect that after the immediate effects of this nutrient pulse have dissipated, the system will recover to the origin (antecedent baseline); however, because of the strong input of nutrients, the lake is less oligotrophic than before. The system has slightly changed, and consequently, the baseline has changed. Thus, there are low expectations that the biological interactions of the system will recover to the same level as before the perturbation (Thayne et al., 2021). Hence, the system might recover to a different level than where the perturbation started, and this happens due to a shift of the baseline in the direction of the regime shift rather than a direct response to the perturbation.

Here, the scale on which the system is assessed plays an important role. Selecting more generic proxies at higher ecological levels (e.g., community) may mask the lack of recovery in specific parts of the ecological network (e.g., population), providing the wrong idea that the system recovered to the same place it was before. This is an important concept when comparing recovery at different moments within the same system (e.g., different years) and between systems (e.g., replicates, treatments, lakes) because the expected trajectory of the recovery will likely be modified by the increased pressure of the driving force.

Earlier in the paper, we pointed out that combining a continuously changing driving force in association with independent perturbations could be a fruitful conceptual framework for bridging theory and real-world application of EWS. However, no complex experimental setup so far has complied with this scenario. Also, no experiment reported the assessment of recovery from natural perturbations (e.g., storms and heatwaves) that may have happened along with the whole-system manipulation experiments. If such events co-occurred with the experimental treatments, they were considered a regular part of the timeseries rather than a potential perturbation (e.g., smoothing these events using moving averages with long rolling windows).

3.1.2. Nature of the forces inducing a regime shift

Above, we described how driving forces and perturbations were experimentally designed. Here we detail the nature of these forces, the observed ecological responses they induced, and what lessons could be learned. We summed up six different types of regime shifts comprising 12 different driving forces. In this session, we are mostly interested in whether the experiments had the capability to bridge theory and application (i.e., not in which results were obtained). For the results of each experiment, see (2) Most reliable combination of proxies and metrics (EWS) and Supplementary material – Metadata.

3.1.2.1. Simple biological setups. In all the simple experiments the driving force pushed the population towards collapse, having extinction as the stated regime shift. This is a fairly understandable regime shift when the system is composed of a single species or based either on a competitive exclusion or a simple predator-prey system. Yet, from an ecological perspective, a limited number of lessons can be learned from extinction as a valid regime shift. The most valuable of them is that if local extinction shows EWS, the collapse of a given taxon within an ecological network may also be anticipated by EWS. This could help identify compensatory dynamics providing resistance to changes at the community level (i.e., sustaining ecological resilience before showing a non-linear response), and its implications are further discussed in section (a) Abundance below.

The perturbations always mimicked mortality events and were mostly applied as serial perturbations. There is no explicit explanation of why this is the case. However, we assume this is due to mortality being a relatively simple experimental manipulation, associated with a straightforward quantification of its direct effects. Nevertheless, other "easy to apply" perturbations like pulses of turbidity, acidification, or a heatwave should be encouraged to expand the range of validation of EWS as a proof-of-concept.

While the nature of the perturbation was always mortality, the nature of the driving forces varied. Veraart et al. (2011) (ID 4) applied mortality pulses against a background of increasing light intensity, resulting in an extinction event by photo-inhibition of a cyanobacterial population (*Aphanizomenon flos-aquae* (L.) Ralfs). Dai et al. (2015) applied mortality pulses while (i) reducing nutrient provisioning for *Saccharomyces cerevisiae* cultures until their collapse by starvation (ID 10) and (ii) increasing mortality rates (population turnover) towards the density-based extracellular digestion threshold, resulting in population metabolic collapse (ID 11). Ghadami et al. (2018) (ID 19) have reproduced the latter experiment with minor modifications, and Sommer et al. (2016) (ID 14) applied mortality pulses in a deteriorating environment of rotifer cultures (*Brachionus calyciflorus*). Those were the five experiments we signalized as a robust proof-of-concept for EWS using an increasing driving force associated with independent perturbations. Food provision and shifts in temperature completed the list of driving forces used in simple experiments.

3.1.2.2. Complex biological setups. Complex biological setups included a wider array of regime shifts compared to the simple ones. We found seven different driving forces applied to nine unique experiments. In a few experiments, perturbations were caused by strong manipulations of the driving force. However, in most cases, the perturbations were independent of the driving force and mimicked mortality events. Below we highlight the individual contributions that experiments provided towards a real-world application of EWS.

The whole-river manipulation studied by Robinson and Uehlinger (2008) (ID 1) focused on the changes and recovery of periphyton assemblages, seston, and macroinvertebrate communities after serial perturbations caused by flood events in the River Spöl, Switzerland. This experiment had no stated driving forces being manipulated and might be a valid example of how additive perturbations can build-up over time, eventually surpassing tipping points, resulting in a regime shift. After a series of flood perturbations, the river shifted from a flood controlled to a flood resilient community. This study was capable of identifying EWS – although not called as such by the authors - before the unfolding regime shift in the third year of the flood program (16 floods in total). It is important to note that the experiment lasted for eight years. It is plausible that pressures at the catchment level may have influenced the river alongside the flood perturbations with such a long timeframe, eventually acting as a hidden driving force. Despite this, the experiment shows that the assessment of serial perturbations can successfully be used to anticipate a regime shift in a real-world situation, as long as the length of the experiment allows multiple pulse perturbations.

The three whole–lake manipulations were sequential experiments done at Paul and Peter Lakes, Wisconsin, USA. The first study started in 2008, and the others that followed carried its legacy effect. The first experiment used multiple proxies to assess EWS in a regime shift from planktivorous to piscivorous fish dominated lake (Carpenter et al. (2008) - ID 3). A significant number of largemouth bass (*Micropterus salmoides*) was added once a year (stepwise increasing driving force), forcing a trophic cascade caused by increased top-down control of the food-web. The regime shift unfolded after the third year of the experiment, resulting in multiple publications (Carpenter et al. (2011), Seekell et al. (2012), Seekell et al. (2013), Batt et al. (2013), Pace et al. (2013), Cline et al. (2014)). Due to the steep increase in the driving force (substantial addition of top predator fish), this was one case where the manipulation triggered an associated perturbation that allowed quantification of recovery rates. This experiment showed that many EWS, as an increase in the variance of grazers and autocorrelation of phytoplankton, can be observed in complex and natural networks with multiple ecological interactions. Yet, the experiment relied on an in-depth understanding of the unfolding processes forcing a regime shift that possibly facilitated the search for EWS. The first follow-up experiment (ID 17 - Wilkinson et al. (2018)) assessed the presence of EWS preceding the occurrence of phytoplankton blooms, using phytoplankton pigments and oxygen as proxies of a regime shift. For that, Peter and a third lake received slow nutrient additions (driving force) over the summer for three consecutive years. The driving force continuously increased over time, and its regime did not trigger a perturbation. This is the only complex biological setup that was not subjected to any kind of induced perturbation, having an increase in the driving force as the sole factor responsible for a regime shift (algal bloom). In the last experiment of the series, Butitta et al. (2017) (ID 15) induced a cyanobacterial bloom in Peter lake by applying a single nutrient pulse (approximately 2.5-times stronger in a third of the time compared to the previous study). Here, the driving force's abrupt regime triggered a perturbation, bringing the system out of equilibrium. A recurrent assumption for the application of EWS is the existence of alternative states (Scheffer et al., 2009); however, in these two experiments, the stated regime shift was an algal bloom. Blooms are often seen as a transient state of the system rather than a permanent regime shift since they often lack proper feedback loops to stabilize the system in the "bloom state" (but see Scheffer et al. (1997), where *Planktothtrix agardhii* blooms create the dark underwater conditions that promoted the continuation of these blooms, excluding other phytoplankton taxa that are less efficient in light harvesting). Still, these last two whole lake eutrophication experiments indicated that EWS can be more versatile than predicted by theory, and indeed observable also when approaching transient regime shifts. EWS were already suggested to anticipate non-catastrophic shifts yet under specific conditions (Dakos et al. (2015); Kéfi et al. (2013)).

We found four coastal ecosystem manipulations. All of them used mortality as a perturbation and were steady-state systems with no increase in the driving force over time; however, the nature of the driving forces and regime shifts differed between the experiments. Two experiments induced a single mortality event (perturbation) of seagrass (*Zostera sp.*) and compared recovery either under different nutrient levels (ID 8 - Soissons et al. (2014)) or different tide levels as the driving force (ID 18 - El-Hacen et al. (2018)). While comparing multiple recovery patterns within the same timeseries is the most straightforward way of inferring levels of resilience, these two experiments mimic the critical slowing down – a temporal phenomenon – using spatial replicates under different pressures of the driving force. The other two coastal manipulations forced a canopy/turf dominated regime shift either by clipping (ID 9 - Benedetti-Cecchi et al. (2015)) or removing canopy cover in different levels of treatment (ID 16 - Rindi et al. (2017); Rindi et al. (2018)). The canopy/turf regime shift is driven by a similar process observed in the macrophyte/phytoplankton alternative states in shallow lakes theory (Scheffer, 2004), showing that analog ecological processes can be observed in different ecosystems. Despite having the driving force as a steady-state variable, all four experiments had a significant period of time between setting the treatment levels and maintaining it during the experiment (e.g., setting canopy cover in 2013 but next time adjusting the level of cover only again in 2014). Thus, these experiments are relevant examples where the length of the experiment might itself act as a composite of driving forces that encompasses many different ecosystem stressors that are not fully controlled throughout the experiment (e.g., water table, nutrient oscillations, seasonality, re-growth).

The last paper is the only complex biological setup that was not a real-system manipulation. Sirota et al. (2013) used the pitcher plant (*Sarracenia purpurea)* – a plant that forms a miniature of an aquatic ecosystem (microcosm) in its lumen with a detritus-based food web composed of five well-determined trophic levels (ID 7) - to anticipate proximity to a regime shift from an aerobic to an anaerobic system. For that, organic matter (driving force) was added to the lumen at four different rates, forcing the system towards the tipping point. The addition of organic matter happened as strong daily pulses, pushing the system out of equilibrium. Thus, this is another example where the driving force can be used to inducing a perturbation. What is peculiar in this study is that all the proxies were based on abiotic parameters, pointing towards the possibility of using physicochemical parameters to assess EWS in shifts in system metabolism.

The various combinations of experimental designs we found on the 19 unique experiments indicate a wide range of scenarios and situations where EWS may potentially be observed. Yet, the primary focus of these experiments was not to tackled inherent issues encountered when applying EWS to non-experimental data (i.e., lack of cohesion between EWS, false results, representative proxies) or producing a mechanistic understanding of how EWS can be optimized for real-world applications. The primary objective of these experiments was to produce a supportive proof-of-concept for the empirical existence of EWS in living systems. This is a paramount milestone that EWS had to achieve before stepping into the future developments needed to comply with the diversity and complexity of challenges posed by the natural systems.

3.2. Most reliable combination of proxies and metrics (EWS)

Our analysis covered 119 quantifications of EWS, of which 35 were from simple biological setups, and 84 were from complex biological setups. Simple biological setups had an overall success rate of 66%, decreasing to 58% in complex biological setups. However, our assessment also showed that some combinations of proxies and metrics are markedly used more often and had a higher success rate than others. We compiled as many as 13 proxies and 17 metrics used to identify EWS of regime shift, totaling 221 possible combinations. The combination of 4 metrics (autocorrelation, recovery, variance, and shape of the distribution) with 4 proxies (abundance, Chlorophyll-a, Phycocyanin, or dissolved oxygen) accounted for 72% of all the quantifications of EWS. For these cases, the overall success rate was about 70%, independently of the biological complexity (Table 2). Below we scrutinize the success rates of these most common combinations, presenting the critical mechanistic aspects of its application. For the other 28% of the results, we mostly compiled too few results (or none) on the same "proxy vs. metric" combination to allow further discussion. For more information on the performance of these combinations, see Supplementary material - All Proxy vs. Metric Combinations. As a reminder, we cannot discard the possibility – as is customary in the way we all publish focusing on positive results only - that proxies and metrics that did not show promising results were not included in the original publications. We call the attention that this could pose a publication bias that could affect the numerical results presented here (while not changing the mechanistic reasoning underlying it).

Table 2
Summary results of proxy vs. metric combinations (EWS) obtained from the literature review of 24 aquatic experiments from early-2020 and earlier." +" is the number of results originally described as positive, "n" the total number of results; "SR" success rate; "partial" refers to the partial sum of all proxies for a given metric according to the biological complexity of the experiment. "Total" refers to the overall results considering all the types of proxies and experiments.

	Autocorrelation			Shape of the Distribution			Recovery			Variance			All metrics		
	+	n	SR	+	n	SR	+	n	SR	+	n	SR	+	n	SR
Simple															
Abundance	7	9	78%	1	4	25%	5	7	71%	9	11	82%	22	31	71%
Chlorophyll-a	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Oxygen	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Phycocyanin	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Partial	*7*	*9*	78%	*1*	*4*	25%	*5*	*7*	71%	*9*	*11*	82%	22	31	71%
Complex															
Abundance	2	3	67%	2	2	100%	3	5	60%	4	4	100%	11	14	79%
Chlorophyll-a	5	6	83%	1	1	100%	3	3	100%	5	7	71%	14	17	82%
Oxygen	4	6	67%	0	2	0%	0	1	0%	5	7	71%	9	16	56%
Phycocyanin	2	4	50%	-	-	-	-	-	-	3	4	75%	5	8	63%
Partial	*13*	*19*	68%	*3*	*5*	60%	*6*	*9*	67%	*17*	*22*	77%	39	55	71%
Total	*20*	*28*	71%	*4*	*9*	44%	*11*	*16*	69%	*26*	*33*	79%	*61*	*86*	71%

3.2.1. *Abundance*

The proxy "abundance" was used 31 times in simple experiments and 14 times in complex experiments. Abundance was either expressed as a concentration (e.g., cells/ml) or a count of organisms, and together they sum 52% of all the EWS quantifications scrutinized here. For simple experiments, abundance showed a consistent success rate using autocorrelation (78%), recovery (71%), and variance (82%), always having extinction as a regime shift. The shape of distribution was a poor EWS (25% success rate, n=4), raising a concern on how reliable it can be for the real world since it failed to prove its concept even in simple and highly controlled biological setups. In complex experiments, the success rates were similar to those in the simple experiments, but fewer quantifications were recorded. All the quantifications using variance (n=4) and the shape of the distribution (n=2) succeeded in finding EWS. Autocorrelation was successful in 2 out of 3 quantifications and recovery rate in 3 out of 5.

The high success rate of abundance as a proxy indicates that EWS of regime shifts can most likely be observed in any individual population, either as a simple experiment or as part of a complex ecological network. A clear asset of using abundance in simple experiments is that the population itself is the core of any observable effect caused by the driving force or perturbation. This excludes the possibility of selecting a misleading proxy and likely contributed to higher success rates. In complex networks, however, EWS of regime shift in a collapsing population may not necessarily function as an EWS of the ecosystem as a whole. It must be considered that the decline of one ecological group might be compensated by another, avoiding the occurrence of significant effects at the ecosystem level. Pragmatically, this means that not all the populations will present a valid EWS for the whole ecosystem, and even if they do, they are likely to have different timing. Hence, assessing abundance as an EWS based on a single population is prone to produce false positives. The opposite is also true; the absence of EWS in a given population may not mean that the ecosystem is not threatened by a regime shift (false negative). In experiments where the system is pushed towards a rather identifiable regime shift, false positives and negatives are easier to control. However, for field data, whole-system manipulations, or even large scale and long-term mesocosm experiments, this context might not hold. In those cases, the timing of the regime shift is unknown and previous knowledge on the ongoing processes forcing the shift becomes paramount for defining the most appropriate sentinel groups to be monitored. Using abundance-based proxies relies on a basic understanding of the structure of the ecological network. A point of future investigations would be where to look for such sentinel groups within an ecological network (e.g., Kuiper et al. (2015)) and on how representative processes derived from experiments would be when extrapolating them to field-data. Unfortunately, the experimental data compiled here are insufficient to make objective inferences on the subject.

3.2.2. *Chlorophyll-a and phycocyanin*

The second most used proxy was Chlorophyll-*a*. It was used 17 times as a proxy for EWS in 3 unique experimental setups, always in whole-system manipulations. Within all the four most-used metrics, Chlorophyll-a provided EWS 14 out of 17 times, with success rates always higher than 70% (Table 2). Phycocyanin was used four times with autocorrelation (Success rate = 50%) and another four times with variance (Success rate = 75%), all in a single experiment identifying algal blooms over multiple years. No simple biological setup used phytoplankton pigments as a proxy for EWS.

Differently from the abundance of specific groups, phytoplankton pigments encompass many taxa into a single community proxy, diminishing the distance between the proxy's response and the ecosystem-level response. Thus, fluorescence probes could be a reliable approach for anticipating a regime shift when a detailed process-based understanding of the system is incomplete. Another asset of using fluorescence is in situations where high-frequency monitoring of taxonomic (or functional) groups might not be a feasible option (e.g., the study area is far from the lab, and frequent sampling is not possible). However, it is important to keep in mind that the signals from fluorescent probes are not straightforward to interpret because of their dependence on environmental factors (e.g., Watras et al. (2017)). Also, phytoplankton blooms are known to be transient in most cases. In other words, they may indicate momentary and non-persistent shifts rather than a change to another stable state (e.g., turbid state). This means that for long-term monitoring of ecological resilience, phytoplankton pigments are prone to produce false positives (EWS but no regime shift), unless transient shifts are also of interest. Furthermore, it is expected that monitoring proxies at the community level reduce the time between spotting EWS and the actual occurrence of a regime shift. Community responses tend to resist changes due to compensatory processes at organismal and population levels within the ecological network (Connell and Ghedini, 2015), which may delay the observation of EWS at a community-level. Independently, the success rate of Chlorophyll-a indicates that EWS of regime shift might also be observable in transient regime shifts when using autocorrelation, variance, or recovery from perturbation - complementarily to its observation linked to alternative stable states. For phycocyanin, the same mechanistic rationale applies, but further studies would be required to consolidate its usefulness as a proxy for EWS.

3.2.3. *Dissolved oxygen*

Dissolved oxygen (DO) – with the same potential for high-frequency monitoring as pigments - was used 16 times as a proxy for EWS in three whole system manipulations. It showed potential for observing EWS using autocorrelation and variance as a proxy, where success rates were 67% and 71%, respectively. The shape of distribution was used twice and recovery from perturbation once, both failing to report EWS in well-stated regime shifts. DO showed the lowest success rates for all the metrics described here.

The experiments listed in this session used DO as a direct measurement of oxygen and not as a proxy for metabolic processes (e.g., primary production). An inherent issue of oxygen measurements in ecology is that its behavior is at the same time response and explanatory variable. This means that oxygen values are strongly determined by biological processes like photosynthesis and respiration, at the same time that oxygen partially determines these biological processes. Thus, DO measurements are entangled between being the cause and the effect of ecological processes, creating a circular causality that may complicate its interpretation as an EWS. On top of that, DO can also be highly modulated by external atmospheric forces not related to in-system processes (e.g., temperature, wind, pressure) that might demand specific cautions when using oxygen as a proxy for EWS. Yet, more studies are needed to clarify why the success rate of DO is lower than the other proxies.

In aquatic sciences, DO has long been used as a proxy to understand shifts in biogeochemical processes like productivity, carbon flux, and other ecosystem metabolism metrics. We found a single whole-lake manipulation assessing gross primary production, net ecosystem production, and respiration as an oxygen-based proxy for EWS. However, it failed in identifying a well-stated regime shift (Batt et al., 2013). Thus, we have no good understanding of whether lake metabolism can be used as a proxy for EWS in near-natural systems.

The success rates of the above-mentioned EWS in documented regime shifts were mostly above 70%. We consider this a high value that should encourage its usage in future studies, either in further complex biological setups or even field campaigns. However, no "silver bullet" is at hand that is capable of translating complex ecosystem responses into a unitary measurement of EWS, and so, its interpretation demands caution. For the application of EWS in real-world situations, it might be crucial to compile multiple EWS into multivariate analysis capable of producing a comprehensive status of the ecosystem as a whole (Clements and Ozgul, 2018), and so reducing potential false results.

3.3. Sampling effect on the success rate of EWS

We investigated whether the type of the experiment, its length, number of sampling points, and the frequency of sampling influenced the success rate of EWS described above. However, none of these parameters showed to be paramount for a successful assessment of EWS in experiments. Our models returned no significance of any of these parameters to the output of the logistic regression model, and both final models had a low and non-significant pseudo-R^2 (Recovery R^2=0.19 p-value=0.28; Other metrics R^2=0.04 p-value=0.18 – See Supplementary material – Sampling Rate).

While previous works using model simulations suggested that EWS are more reliable when sampling points are abundant (e.g., Dakos et al. (2012), Clements et al. (2015)), our assessment of the experiments does not support this statement. Experimental systems with similar sampling regimes presented distinct success rates, and different proxies within the same experiment also showed divergent responses. We relate this to the nature of the proxies assessed and interpret it as an indication that no sampling regime may converge towards a pragmatic universal sampling threshold. Inherent characteristics of the studied system (e.g., process rates, population turnover) might be more critical to determine the frequency and number of sampling points than a mere statistical sufficiency. A novel study that combined these two aspects suggested a minimum number of generations to be sampled to increase the reliability of EWS (Arkilanian et al., 2020). With future experiments focusing on the methodological reliability of EWS, we could elucidate for instance, if the sampling regime would have to be tailored based on the expected processes that are pushing the system towards the regime shift.

While so far experiments cannot suggest sampling rates across different systems, we observed that the recovery from perturbation was always inconclusive when sampling was infrequent (less than five sampling points - Supplementary material – Sampling Rate). The recovery from perturbations is mostly related to lifecycle processes (e.g., the re-establishment of the carrying capacity or community turnover) and its assessments based on a low number of sampling points may miss the recovery trajectory, causing inconclusive results (e.g., the system has already recovered for a while before being sampled again). Thus, while not providing a benchmark sampling rate, it is reasonable to say that the assessment of recovery requires a sampling frequency that provides a high resolution of the recovery trajectory whenever extensive and high-frequency sampling is not possible.

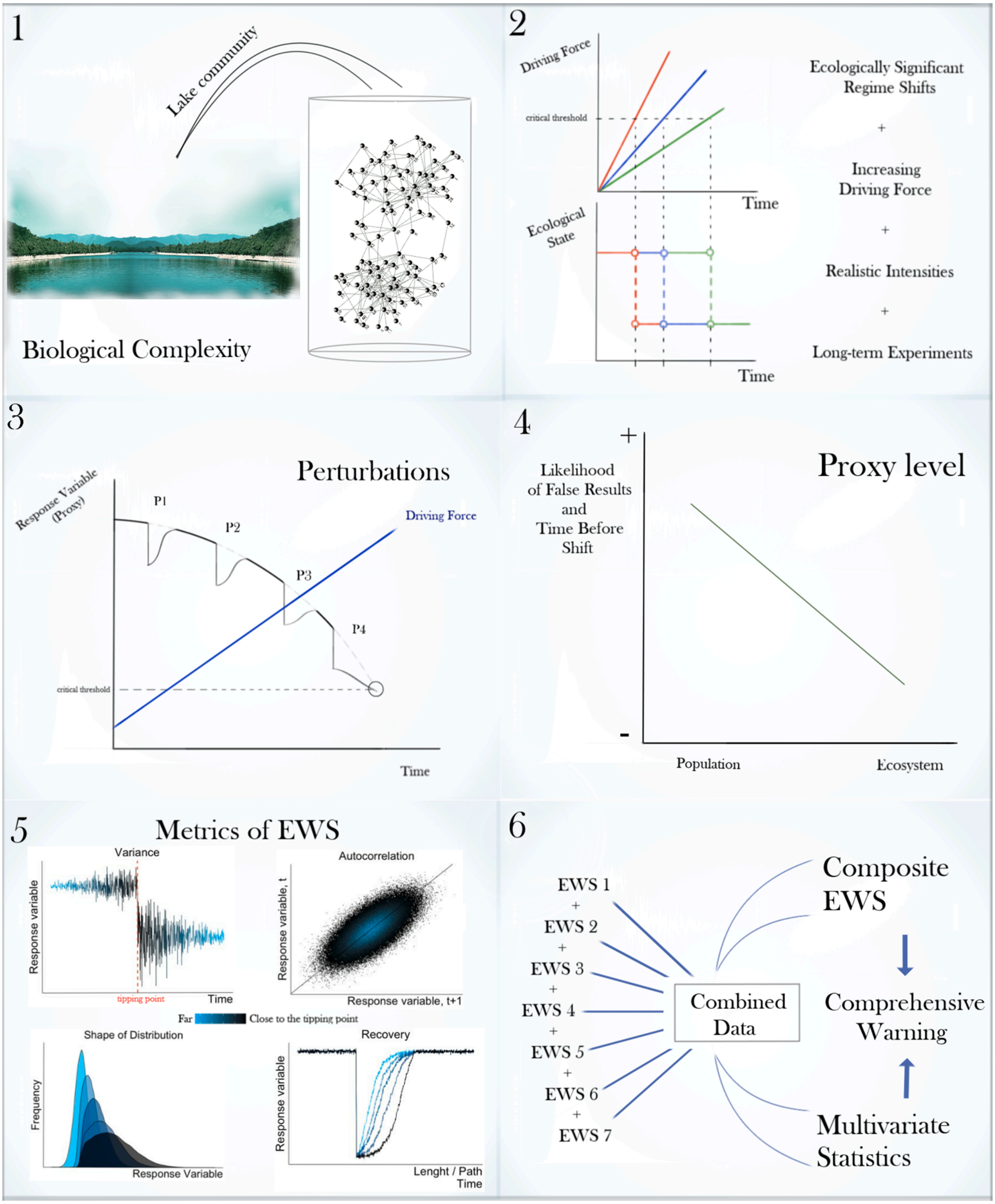
4. Next generation of EWS experiments

Experiments are often a trade-off between our capability for a mechanistic understanding of the system and its proximity to reality. While single-species setups and whole-system manipulations represent the far opposite of this spectrum, both provided significant contributions for the consolidation of EWS. However, our results demonstrate that experiments with near-natural and realistic designs that mimic real-world situations are still rare. The majority of the experiments focused on proving the existence of EWS rather than addressing specific needs for the use of EWS as a management tool. To further support the implementation of EWS at the monitoring level, experiments would have to foresee and address the caveats observed in real-world situations. These are (i) our capability to select measurable variables that are relevant for whole-system analysis (Burthe et al., 2016), (ii) a lack of coherence between different EWS indicators - (e.g., when different metrics applied with the same response variable present divergent results - see Gsell et al. (2016), Burthe et al. (2016)), (iii) the requirement for ecosystem-specific knowledge of transition-generating mechanisms (Gsell et al. (2016), Spears et al. (2017)), and (iv) data acquisition issues. So far, there is no solid applicable experimental framework capable of steering the field of EWS towards real-world applications. A fundamental conceptual starting point is that EWS experiments are resilience experiments. A successful resilience experiment must answer the following questions: (i) resilience of what?, (ii) to what? and (iii) compared to when? (Carpenter et al., 2014; Hodgson et al., 2015). Below, we selected six aspects we considered fundamental for bringing EWS experiments closer to real-world situations, promoting the future application of EWS. The numbers of the headings represent the panels in Figure 1.

4.1. Biological complexity

Most EWS experiments so far were simple biological setups, hampering the understanding of where to look for EWS in nature. Future experiments would have to consider (near-) natural communities to mimic (near-) natural ecological processes. Whole-system manipulations would be the closest scenario for bridging EWS observed in experiments to the natural environments. They provide insightful details on processes that could cascade into a regime shift and provide realistic quantification for those processes (Malley and Mills, 1992), which is paramount for making inferences about resilience (Pimm et al., 2019). Yet, whole-system manipulations are often difficult to implement for various reasons and remain relatively rare in ecology.

A feasible alternative to whole-system manipulations is the use of mesocosms filled with local lake water. Mesocosms provide an intermediate step between simple and the most complex biological setups and can potentially mimic complex ecological networks (Cadotte et al., 2005; Stewart et al., 2013). It is known that mesocosms may lack fundamental aspects of resilience - such as metacommunity structure (e.



(caption on next page)

Fig. 1. Six steps to consider for bringing EWS experiments closer to a real-world situation. The number of the panel represents the heading on the main text. (1) The use of whole-lake manipulations or mesocosms built using natural lake communities helps to incorporate biological complexity into the observed responses. Complex ecological networks have inherent proprieties that support resilience and resistance to changes, which is difficult to mimic when using artificial ecological networks. (2) Experiments should aim to slowly increase the driving force over time until reaching the tipping point, instead of using steady-state treatments. Continuous increase in the driving force changes the shape of the basin of attraction during the experiment, pushing the system closer to the tipping point over time. This is the situation that comes closer to real-world situations. Whenever possible, use different rates of increase as treatments. This would allow us to understand how the intensity of the driving force can modulate the timing of the regime shift. (3) Assessing serial pulse perturbations within the same systems over time adds a simple and easily quantifiable metric of recovery that can provide insights on levels of ecological resilience. Also, the knowledge obtained from recovery from serial perturbations can more easily be transferable to real-world applications (e.g., recovery from storms as an indicator of lake resilience). (4) Responses at a population level may not upscale to a community or ecosystem-level. Thus, the lower the ecological level of the proxy, the higher is the risk of obtaining a false result (observing an EWS but no regime shift or vice versa). At the same time, populations respond more promptly to disturbances and are expected to deliver earlier EWS compared to other ecological levels. The inherent a priori knowledge of an experimental system may help to select lower ecological levels with high ecosystem relevance (e.g., engineering species), reducing the likelihood of false results while providing longer intervals between observing an EWS and the actual shift. Common examples of (i) population-level proxies: species abundance or biomass; (ii) community: phytoplankton fluorescence or abundance of grazers (iii) ecosystem: lake respiration, turbidity, or oxygen saturation. (5) Metrics of variance, autocorrelation, shape of the distribution, and recovery from perturbation are the most reliable metrics to be applied in future studies aiming to bridge the gap between theory and real-world application of EWS. The increase in the driving force with time – panel 2 - pushes the system towards the tipping point. When getting closer to the tipping point, variance, autocorrelation, and skewness are expected to increase, while recovery from perturbation is expected to get slower. These four leading indicators showed a high success rate in aquatic EWS experiments. (6) Individual EWS are unlikely to produce a comprehensive overview of what is happening in a complex system. However, the different EWS can be used to form a composite signal or become part of a multivariate analysis to produce a comprehensive ecosystem-level warning. Composites are based on normalization and standardization of multiple EWS, while multivariate statistics may integrate cross-correlation between EWS to produce a comprehensive warning.

g., Bengtsson (2002); Isaac et al. (2018); Oliver et al. (2015); Peterson et al. (1998)) – and the process rates calculated from it are unlikely to be fully representative of natural environments (e.g., Malley and Mills (1992)). Still, mesocosms can help to identify the nature and order of unfolding processes triggering a regime shift. Compared to simple experimental setups or artificially constructed networks, near-natural mesocosms can also incorporate important aspects of resilience as functional redundancy (see Walker (1992)) and compensatory processes (see Brown et al. (2016); Connell and Ghedini (2015)). Thus, they become more insightful to address complex ecosystem responses when whole-system manipulations are not possible and remain an important tool if we intend to understand where to look for EWS in real-world situations.

4.2. Perturbations

Experiments showed that assessing recovery from perturbations is a reliable and straightforward measurement of resilience. Having perturbations that simulate natural processes incorporated into complex biological setups can help developing valuable methods for assessing resilience in natural conditions (e.g., storm-alike events). Perturbations are inherent events of natural systems and can be a valuable opportunity for translating recovery in experiments to the real-world assessment of resilience. Also, assessing recovery from perturbations demands only a fraction of the data needed for other EWS and can help when data acquisition is an issue (e.g., long-term data is lacking, or monitoring campaigns cannot operate continuously). A consistent challenge is that assessing recovery from perturbations may require a higher sampling frequency before, during, and after the perturbation. EWS literature lacks the information on how high the sampling frequency should be for producing a reliable assessment of recovery; therefore, future experiments should apply a sampling rate that matches the scale of the perturbation and the rate of turnover of the selected proxies while minimally complying with the Nyquist frequency to avoid aliasing effect (especially when assessing recovery). For example, the assessment of a storm whose direct effects last for a couple of days will demand proxies with a high population turnover (e.g., phyto- and zooplankton), while assessment of a hurricane whose direct effects may last months may allow proxies with a lower population turnover (e.g., macrophytes or fish). To capture the fast response of plankton to the storm, sampling rates will have to be higher than the one required for capturing the response of macrophytes to the hurricane.

Perturbations are known to be capable of muffling EWS by altering autocorrelation, variability, and modifying signal-to-noise ratio (e.g., Clements et al. (2015); Garcia-Gudino et al. (2017)). However, and despite it, mimicking them in experiments should not be seen as a complicating factor for analysis but as an opportunity for developing novel toolkits capable of incorporating them into the analysis. Inducing serial perturbations of different magnitudes and natures throughout the experiment may help develop tools that are more easily transferable to real-world applications while still minimizing data acquisition issues.

4.3. Slow increase in the driving force

A fundamental caveat to be overcome by experiments relates to the nature, intensity, and regime of the driving force applied to the systems. First, experiments should use driving forces that are commonly acknowledged as real-world environmental pressures, forcing the system towards a regime-shifts of ecological relevance. Most of the experiments done so far were composed of 1-2 species artificially pushed towards extinction, which does not facilitate real-world applications of EWS at the ecosystem level. It is easy to cause local extinction in experiments, but in real-world conditions local extinction is rare (see Figueiredo et al. (2019) and Harrison (1991)). Second, driving forces have to be scaled to rates that are realistic to real-world processes (see Korell et al. (2020)). Driving forces in experimental data are often applied at a much steeper gradient than in the real world (Alborzi et al., 2018; Benedetti-Cecchi et al., 2015; Korell et al., 2020) or conversely, are assumed to be in a steady-state. While this strategy facilitates the assessment of the experiment, it does not facilitate the extrapolations of analytical tools to real-world situations. Driving forces would ideally change gradually over time, slowly eroding its basin of attraction, allowing time for the system to respond. For that, resilience experiments with complex biological assemblages would have to be designed as long-term experiments, lasting several weeks to even months or years, so they can better mimic the temporal scale of changes from a real-world situation. This is also the scenario where EWS are expected to be more easily observed (e.g., Arkilanian et al. (2020); Clements and Ozgul (2016b); Dakos et al. (2015)). Lastly, applying different rates of increasing driving force would help us understand how the intensity of the driving force can modulate the timing of the regime shift.

4.4. Defining proxies

Key information experiments may provide us with is where best to look for EWS of regime shifts. Unfortunately, EWS experiments were not designed or assessed to explore this question. As discussed before, proxies can be defined at many different levels of ecological organization (i.e., population, community, and ecosystem-level). The proxy's ecological level is expected to affect how long before the actual regime shift an EWS is observed and its potential for presenting false results (see session (2) Most reliable combination of proxies and metrics (EWS)). The closer a proxy is to a population level, the further it is from representing ecosystem-level responses (Cottingham and Carpenter, 1998). On the other hand, populations respond faster to environmental changes than communities or ecosystems, and so are expected to show more prompt response to environmental changes. While this concept is broadly consolidated in ecology (e.g., Brown et al. (2016); Connell and Ghedini (2015); Klug et al. (2000); Steiner et al. (2006)), we lack specific data confirming such a pattern when applying EWS.

Process-based experiments can couple the responses of population-, trait-, community- and ecosystem-based proxies for the same ecological process. In this way, experiments can help translate how signals from populations can be reliably upscaled to community and ecosystem levels. Such an approach is powerful, not only in providing information on which groups are significant to monitor but when in time EWS are expected to scale up to higher ecological levels (successional patterns towards the tipping point). Abundance-based proxies were applied in most of the aquatic experiments performed so far and showed promising success rates. Overall, most of the long-term ecological monitoring data for aquatic systems are also population-based (e.g., taxonomic lists). Setting up experiments capable of foreseeing the use of this already existent monitoring data would be a valuable step to bridge the gap between theory and application.

4.5. Metrics

Although many different metrics have been tested in aquatic experiments, only four of them accumulated enough data to support further use in real systems. So far, most experiments obtained reliable results using autocorrelation, variance, recovery, and sometimes shape of the distribution for identifying EWS in well-stated regime shifts. Although, across experiments, no metric showed to work 100% of the time. Also, different metrics like autocorrelation and variance are reported sometimes to produce conflicting results even when assessing the same proxy under the same conditions (see Gsell et al. (2016), and the Supplementary material - Metadata). Such a caveat undermines our capability to understand if the EWS is in fact happening or if the response is a mere statistical artifact. Future experiments can potentialize their application to real-world situations by varying sampling frequencies and using timeseries with different levels of noise. Assessing all the main leading indicators of regime shift within the same proxy may clarify why we see antagonistic results (e.g., if due to low signal-to-noise ratio) as well as investigate ways to minimize it (e.g., changes in the frequency and length of sampling).

4.6. Composites and multivariate analysis

Experiments showed that a single EWS - the combination of a proxy and a metric - will not be capable of addressing the full complexity embedded in approaching a regime shift. Different EWS could be most useful as a partial indication that changes are about to happen in the studied system (Clements and Ozgul (2018); Eason et al. (2016); Lindegren et al. (2012)). A vital aspect that can be learned from experiments is how to merge responses from different EWS into a comprehensive composite of early warning signals (e.g., Clements and Ozgul (2016a); Drake and Griffen (2010)). Composites are based on a relatively simple normalization and standardization of multiple EWS that create an EWs index. Another possibility is the use of multivariate analysis – an analysis capable of collapsing the behavior of multiple variables of a complex system into an index that captures the system dynamics over time, space, or both (Fath et al., 2003) - to help us translating individual EWS into significant ecosystem-level responses.

Recent studies have reported promising results towards the observation of *a posteriori* regime shifts in simulated data (Eason et al. (2013); Karunanithi et al. (2008)), paleoecological data (Eason et al. (2016); Spanbauer et al. (2014)), and field data (Eason et al. (2013); Eason et al. (2016); Karunanithi et al. (2008); Sundstrom et al. (2017)) using a multivariate temporal data analysis framework (e.g., information theory). Nevertheless, the use of these tools to assess multiple EWS *a priori* is understudied. Near-natural experiments focused on process-based assessments would help shaping future multivariate frameworks applicable to monitoring and management on the grounds of EWS.

5. Conclusion

The experiments presented here should suffice as empirical proof of the existence of EWS in aquatic systems. Success rates of the most used EWS (Proxies: abundance, Chlorophyll-a, Phycocyanin, and dissolved oxygen; Metrics: autocorrelation, variance, recovery, and shape of the distribution) were often higher than 70%. Yet, no EWS showed to be 100% reliable, and their use demands caution. Furthermore, experiments of different natures have shown us that the use of EWS can be more versatile than expected, going from monitoring of extinction of single populations to anticipation of transient regime shifts. However, experiments were few in number and often not designed to tackle issues encountered in real-world situations. Consequently, they lack a deep mechanistic understanding of why, when, and how an EWS is observed or not. To bridge the gap between theory and application, experiments will have to come closer to real-world conditions whilst at the same time support a mechanistic understanding of why EWS may succeed or fail to anticipate a regime shift. For that, we proposed a combined experimental scenario that could enhance the capabilities of EWS to go beyond the stage of proof-of-concept, foreseeing situations compatible with real-world application, for the benefit of ecosystem management and nature preservation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors thank V. Dakos for his useful comments on earlier versions of this manuscript. J.A.A.S and J.P.M were funded by the MANTEL ITN (Management of climatic extreme events in lakes and reservoirs for the protection of ecosystem services) through the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska Curie grant agreement no. 722518.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ecocom.2021.100944.

References

Alborzi, A., Mirchi, A., Moftakhari, H., Mallakpour, I., Alian, S., Nazemi, A., Hassanzadeh, E., Mazdiyasni, O., Ashraf, S., Madani, K., Norouzi, H., Azarderakhsh, M., Mehran, A., Sadegh, M., Castelletti, A., AghaKouchak, A., 2018. Climate-informed environmental inflows to revive a drying lake facing meteorological and anthropogenic droughts. Environ. Res. Lett. 13.

Arkilanian, A.A., Clements, C.F., Ozgul, A., Baruah, G., 2020. Effect of time series length and resolution on abundance- and trait-based early warning signals of population declines. Ecology 101, e03040.

Batt, R.D., Carpenter, S.R., Cole, J.J., Pace, M.L., Johnson, R.A., 2013. Changes in ecosystem resilience detected in automated measures of ecosystem metabolism during a whole-lake manipulation. Proc. Natl. Acad. Sci. U. S. A. 110, 17398–17403.

Benedetti-Cecchi, L., Tamburello, L., Maggi, E., Bulleri, F., 2015. Experimental Perturbations Modify the Performance of Early Warning Indicators of Regime Shift. Curr. Biol. 25, 1867–1872.

Bengtsson, J., 2002. Disturbance and resilience in soil animal communities. Eur. J. Soil Biol. 38, 119–125.

Bestelmeyer, B.T., Ellison, A.M., Fraser, W.R., Gorman, K.B., Holbrook, S.J., Laney, C.M., Ohman, M.D., Peters, D.P.C., Pillsbury, F.C., Rassweiler, A., Schmitt, R.J., Sharma, S., 2011. Analysis of abrupt transitions in ecological systems. Ecosphere 2.
Boettiger, C., Hastings, A., 2012. Early warning signals and the prosecutor's fallacy. Proc. Biol. Sci. 279, 4734–4739.
Boettiger, C., Ross, N., Hastings, A., 2013. Early warning signals: the charted and uncharted territories. Theor. Ecol.-Neth. 6, 255–264.
Brock, W.A., Carpenter, S.R., 2010. Interacting regime shifts in ecosystems: implication for early warnings. Ecol. Monogr. 80, 353–367.
Brown, B.L., Downing, A.L., Leibold, M.A., 2016. Compensatory dynamics stabilize aggregate community properties in response to multiple types of perturbations. Ecology 97, 2021–2033.
Burthe, S.J., Henrys, P.A., Mackay, E.B., Spears, B.M., Campbell, R., Carvalho, L., Dudley, B., Gunn, I.D.M., Johns, D.G., Maberly, S.C., May, L., Newell, M.A., Wanless, S., Winfield, I.J., Thackeray, S.J., Daunt, F., Allen, C., 2016. Do early warning indicators consistently predict nonlinear change in long-term ecological data? J. Appl. Ecol. 53, 666–676.
Butitta, V.L., Carpenter, S.R., Loken, L.C., Pace, M.L., Stanley, E.H., 2017. Spatial early warning signals in a lake manipulation. Ecosphere 8.
Cadotte, M.W., Drake, J.A., Fukami, T., 2005. Constructing nature: laboratory models as necessary tools for investigating complex ecological communities. Advances in Ecological Research. Academic Press, pp. 333–353.
Carpenter, S., Walker, B., Anderies, J.M., Abel, N., 2014. From metaphor to measurement: resilience of what to what? Ecosystems 4, 765–781.
Carpenter, S.R., Brock, W.A., Cole, J.J., Kitchell, J.F., Pace, M.L., 2008. Leading indicators of trophic cascades. Ecol. Lett. 11, 128–138.
Carpenter, S.R., Cole, J.J., Pace, M.L., Batt, R., Brock, W.A., Cline, T., Coloso, J., Hodgson, J.R., Kitchell, J.F., Seekell, D.A., Smith, L., Weidel, B., 2011. Early warnings of regime shifts: a whole-ecosystem experiment. Science 332, 1079–1082.
Clements, C.F., Drake, J.M., Griffiths, J.I., Ozgul, A., 2015. Factors influencing the detectability of early warning signals of population collapse. Am. Nat. 186, 50–58.
Clements, C.F., Ozgul, A., 2016a. Including trait-based early warning signals helps predict population collapse. Nat. Commun. 7, 10984.
Clements, C.F., Ozgul, A., 2016b. Rate of forcing and the forecastability of critical transitions. Ecol. Evol. 6, 7787–7793.
Clements, C.F., Ozgul, A., 2018. Indicators of transitions in biological systems. Ecol. Lett. 21, 905–919.
Cline, T.J., Seekell, D.A., Carpenter, S.R., Pace, M.L., Hodgson, J.R., Kitchell, J.F., Weidel, B.C., 2014. Early warnings of regime shifts: evaluation of spatial indicators from a whole-ecosystem experiment. Ecosphere, 5.
Connell, S.D., Ghedini, G., 2015. Resisting regime-shifts: the stabilising effect of compensatory processes. Trends Ecol. Evol. 30, 513–515.
Cottingham, K.L., Carpenter, S.R., 1998. Population, community, and ecosystem variates as ecological indicators: Phytoplankton responses to whole-lake enrichment. Ecol. Appl. 8, 508–530.
Dai, L., Korolev, K.S., Gore, J., 2013. Slower recovery in space before collapse of connected populations. Nature 496, 355–358. https://doi.org/10.1038/nature12071.
Dai, L., Korolev, K.S., Gore, J., 2015. Relation between stability and resilience determines the performance of early warning signals under different environmental drivers. Proc. Natl. Acad. Sci. U. S. A. 112, 10056–10061.
Dai, L., Vorselen, D., Korolev, K.S., Gore, J., 2012. Generic indicators for loss of resilience before a tipping point leading to population collapse. Science 336, 1175–1177. https://doi.org/10.1126/science.1219805.
Dakos, V., Carpenter, S.R., Brock, W.A., Ellison, A.M., Guttal, V., Ives, A.R., Kefi, S., Livina, V., Seekell, D.A., van Nes, E.H., Scheffer, M., 2012. Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data. PLoS One 7, e41010.
Dakos, V., Carpenter, S.R., van Nes, E.H., Scheffer, M., 2015. Resilience indicators: prospects and limitations for early warnings of regime shifts. Philos. T R Soc. B 370, 20130263.
Dakos, V., Kefi, S., Rietkerk, M., van Nes, E.H., Scheffer, M., 2011. Slowing down in spatially patterned ecosystems at the brink of collapse. Am. Nat. 177, E153–E166.
Drake, J.M., Griffen, B.D., 2010. Early warning signals of extinction in deteriorating environments. Nature 467, 456–459.
Eason, T., Garmestani, A.S., Cabezas, H., 2013. Managing for resilience: early detection of regime shifts in complex systems. Clean Technol. Environ. Policy 16, 773–783.
Eason, T., Garmestani, A.S., Stow, C.A., Rojo, C., Alvarez-Cobelas, M., Cabezas, H., 2016. Managing for resilience: an information theory-based approach to assessing ecosystems. J. Appl. Ecol. 53, 656–665.
El-Hacen, E.M., Bouma, T.J., Fivash, G.S., Sall, A.A., Piersma, T., Olff, H., Govers, L.L., 2018. Evidence for 'critical slowing down' in seagrass: a stress gradient experiment at the southern limit of its range. Sci. Rep. 8, 17263.
Fath, B.D., Cabezas, H., Pawlowski, C.W., 2003. Regime changes in ecological systems: an information theory approach. J. Theoretical Biol. 222, 517–530.
Figueiredo, L., Krauss, J., Steffan-Dewenter, I., Cabral, J.S., 2019. Understanding extinction debts: spatio-temporal scales, mechanisms and a roadmap for future research. Ecography 42, 1973–1990.
Garcia-Gudino, D., Landa, E., Mendoza-Temis, J., Albarado-Ibanez, A., Toledo-Roy, J.C., Morales, I.O., Frank, A., 2017. Enhancement of early warning properties in the Kuramoto model and in an atrial fibrillation model due to an external perturbation of the system. PLoS One 12, e0181953.
Ghadami, A., Gourgou, E., Epureanu, B.I., 2018. Rate of recovery from perturbations as a means to forecast future stability of living systems. Sci. Rep. 8, 9271.
Gsell, A.S., Scharfenberger, U., Ozkundakci, D., Walters, A., Hansson, L.A., Janssen, A.B., Noges, P., Reid, P.C., Schindler, D.E., Van Donk, E., Dakos, V., Adrian, R., 2016. Evaluating early-warning indicators of critical transitions in natural aquatic ecosystems. Proc. Natl. Acad. Sci. U. S. A. 113, E8089–E8095.
Harrison, S., 1991. Local extinction in a metapopulation context - an empirical-evaluation. Biol. J. Linnean Soc. 42, 73–88.
Hodgson, D., McDonald, J.L., Hosken, D.J., 2015. What do you mean, 'resilient'? Trends Ecol. Evol. 30, 503–506.
Holling, C.S., 1973. Resilience and stability of ecological systems. Ann. Rev. Ecol. Systemat. 4, 1–23.
Isaac, N.J.B., Brotherton, P.N.M., Bullock, J.M., Gregory, R.D., Boehning-Gaese, K., Connor, B., Crick, H.Q.P., Freckleton, R.P., Gill, J.A., Hails, R.S., Hartikainen, M., Hester, A.J., Milner-Gulland, E.J., Oliver, T.H., Pearson, R.G., Sutherland, W.J., Thomas, C.D., Travis, J.M.J., Turnbull, L.A., Willis, K., Woodward, G., Mace, G.M., 2018. Defining and delivering resilient ecological networks: nature conservation in England. J. Appl. Ecol. 55, 2537–2543.
Jarvis, L., McCann, K., Tunney, T., Gellner, G., Fryxell, J.M., 2016. Early warning signals detect critical impacts of experimental warming. Ecol Evol 6, 6097–6106. https://doi.org/10.1002/ece3.2339.
Karunanithi, A.T., Cabezas, H., Frieden, B.R., Pawlowski, C.W., 2008. Detection and Assessment of Ecosystem Regime Shifts from Fisher Information. Ecol. Soc. 13.
Kéfi, S., Dakos, V., Scheffer, M., Van Nes, E.H., Rietkerk, M., 2013. Early warning signals also precede non-catastrophic transitions. Oikos 122, 641–648.
Klug, J.L., Fischer, J.M., Ives, A.R., Dennis, B., 2000. Compensatory dynamics in planktonic community responses to pH perturbations. Ecology 81, 387–398.
Korell, L., Auge, H., Chase, J.M., Harpole, S., Knight, T.M., 2020. We need more realistic climate change experiments for understanding ecosystems of the future. Glob Chang. Biol. 26, 325–327.
Krkosek, M., Drake, J.M., 2014. On signals of phase transitions in salmon population dynamics. Proc. Biol. Sci. 281, 20133221.
Kuiper, J.J., van Altena, C., de Ruiter, P.C., van Gerven, L.P.A., Janse, J.H., Mooij, W.M., 2015. Food-web stability signals critical transitions in temperate shallow lakes. Nat. Commun. 6, 7727.
Lindegren, M., Dakos, V., Groger, J.P., Gardmark, A., Kornilovs, G., Otto, S.A., Mollmann, C., 2012. Early detection of ecosystem regime shifts: a multiple method evaluation for management application. PLoS One 7, e38410.
Litzow, M.A., Mueter, F.J., Urban, J.D., 2013. Rising catch variability preceded historical fisheries collapses in Alaska. Ecol. Appl. 23, 1475–1487.
Malley, D.F., Mills, K.H., 1992. Whole-lake experimentation as a tool to assess ecosystem health, response to stress and recovery: the Experimental Lakes Area experience. J. Aquatic Ecosyst. Health 1, 159–174.
Oliver, T.H., Heard, M.S., Isaac, N.J., Roy, D.B., Procter, D., Eigenbrod, F., Freckleton, R., Hector, A., Orme, C.D.L., Petchey, O.L. J.T.i.e., evolution, 2015. Biodiversity and resilience of ecosystem functions. 30, 673-684.
Pace, M.L., Carpenter, S.R., Johnson, R.A., Kurtzweil, J.T., 2013. Zooplankton provide early warnings of a regime shift in a whole lake manipulation. Limnol. Oceanogr. 58, 525–532.
Perretti, C.T., Munch, S.B., 2012. Regime shift indicators fail under noise levels commonly observed in ecological systems. Ecol. Appl. 22, 1772–1779.
Peterson, G., Allen, C.R., Holling, C.S., 1998. Original articles: ecological resilience, biodiversity, and scale. Ecosystems 1, 6–18.
Pimm, S.L., 1984. The complexity and stability of ecosystems. Nature 307, 321–326.
Pimm, S.L., Donohue, I., Montoya, J.M., Loreau, M., 2019. Measuring resilience is essential if we are to understand it. Nat Sustain 2, 895–897.
R Core Team, 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
Rindi, L., Bello, M.D., Dai, L., Gore, J., Benedetti-Cecchi, L., 2017. Direct observation of increasing recovery length before collapse of a marine benthic ecosystem. Nat. Ecol. Evol. 1, 153.
Rindi, L., Dal Bello, M., Benedetti-Cecchi, L., 2018. Experimental evidence of spatial signatures of approaching regime shifts in macroalgal canopies. Ecology 99, 1709–1715.
Ripley, B., Venables, B., Bates, D.M., Hornik, K., Gebhardt, A., Firth, D., Ripley, M.B., 2013. Package 'mass. Cran R 538.
Robinson, C.T., Uehlinger, U., 2008. Experimental floods cause ecosystem regime shift in a regulated river. Ecol. Appl. 18, 511–526.
Scheffer, M., 2004. Ecology of Shallow Lakes. Springer Science & Business Media.
Scheffer, M., Bascompte, J., Brock, W.A., Brovkin, V., Carpenter, S.R., Dakos, V., Held, H., van Nes, E.H., Rietkerk, M., Sugihara, G., 2009. Early-warning signals for critical transitions. Nature 461, 53–59.
Scheffer, M., Carpenter, S.R., Dakos, V., van Nes, E.H., 2015. Generic indicators of ecological resilience: inferring the chance of a critical transition. Ann. Rev. Ecol., Evol. Systemat. 46, 145.
Scheffer, M., Carpenter, S.R., Lenton, T.M., Bascompte, J., Brock, W., Dakos, V., van de Koppel, J., van de Leemput, I.A., Levin, S.A., van Nes, E.H., Pascual, M., Vandermeer, J., 2012. Anticipating critical transitions. Science 338, 344–348.
Scheffer, M., Rinaldi, S., Gragnani, A., Mur, L.R., vanNes, E.H., 1997. On the dominance of filamentous cyanobacteria in shallow, turbid lakes. Ecology 78, 272–282.
Seekell, D.A., Carpenter, S.R., Cline, T.J., Pace, M.L., 2012. Conditional heteroskedasticity forecasts regime shift in a whole-ecosystem experiment. Ecosystems 15, 741–747.
Seekell, D.A., Cline, T.J., Carpenter, S.R., Pace, M.L., 2013. Evidence of alternate attractors from a whole-ecosystem regime shift experiment. Theor. Ecol.-Neth. 6, 385–394.
Sirota, J., Baiser, B., Gotelli, N.J., Ellison, A.M., 2013. Organic-matter loading determines regime shifts and alternative states in an aquatic ecosystem. Proc. Natl. Acad. Sci. U. S. A. 110, 7742–7747.

Soissons, L.M., Han, Q., Li, B., van Katwijk, M.M., Ysebaert, T., Herman, P.M.J., Bouma, T.J., 2014. Cover versus recovery: contrasting responses of two indicators in seagrass beds. Mar. Pollut. Bull. 87, 211–219.

Sommer, S., van Benthem, K.J., Fontaneto, D., Ozgul, A., 2016. Are generic early-warning signals reliable indicators of population collapse in rotifers? Hydrobiologia 796, 111–120.

Spanbauer, T.L., Allen, C.R., Angeler, D.G., Eason, T., Fritz, S.C., Garmestani, A.S., Nash, K.L., Stone, J.R., 2014. Prolonged instability prior to a regime shift. PLoS One 9, e108936.

Spears, B.M., Futter, M.N., Jeppesen, E., Huser, B.J., Ives, S., Davidson, T.A., Adrian, R., Angeler, D.G., Burthe, S.J., Carvalho, L., Daunt, F., Gsell, A.S., Hessen, D.O., Janssen, A.B.G., Mackay, E.B., May, L., Moorhouse, H., Olsen, S., Sondergaard, M., Woods, H., Thackeray, S.J., 2017. Ecological resilience in lakes and the conjunction fallacy. Nat. Ecol. Evol. 1, 1616–1624.

Steiner, C.F., Long, Z.T., Krumins, J.A., Morin, P.J., 2006. Population and community resilience in multitrophic communities. Ecology 87, 996–1007.

Stewart, R.I.A., Dossena, M., Bohan, D.A., Jeppesen, E., Kordas, R.L., Ledger, M.E., Meerhoff, M., Moss, B., Mulder, C., Shurin, J.B., Suttle, B., Thompson, R., Trimmer, M., Woodward, G., 2013. Mesocosm experiments as a tool for ecological climate-change research. in: Woodward, G., O'Gorman, E.J. (Eds.). Global Change in Multispecies Systems: Part 3. Academic Press, pp. 71–181.

Sundstrom, S.M., Eason, T., Nelson, R.J., Angeler, D.G., Barichievy, C., Garmestani, A.S., Graham, N.A., Granholm, D., Gunderson, L., Knutson, M., Nash, K.L., Spanbauer, T., Stow, C.A., Allen, C.R., 2017. Detecting spatial regimes in ecosystems. Ecol. Lett. 20, 19–32.

Thayne, M.W., Kraemer, B.M., Mesman, J.P., Ibelings, B.W., Adrian, R., 2021. Antecedent lake conditions shape resistance and resilience of a shallow lake ecosystem following extreme wind storms. Limnol Oceanogr n/a. https://doi.org/10.1002/lno.11859.

van Nes, E.H., Scheffer, M., 2007. Slow recovery from perturbations as a generic indicator of a nearby catastrophic shift. Am. Nat. 169, 738–747.

Veraart, A.J., Faassen, E.J., Dakos, V., van Nes, E.H., Lurling, M., Scheffer, M., 2011. Recovery rates reflect distance to a tipping point in a living system. Nature 481, 357–359.

Walker, B.H., 1992. Biodiversity and Ecological Redundancy. Conserv. Biol. 6, 18–23.

Watras, C.J., Morrison, K.A., Rubsam, J.L., Hanson, P.C., Watras, A.J., LaLiberte, G.D., Milewski, P., 2017. A temperature compensation method for chlorophyll and phycocyanin fluorescence sensors in freshwater. Limnol. Oceanogr.-Methods 15, 642–652.

Wilkinson, G.M., Carpenter, S.R., Cole, J.J., Pace, M.L., Batt, R.D., Buelo, C.D., Kurtzweil, J.T., 2018. Early warning signals precede cyanobacterial blooms in multiple whole-lake experiments. Ecol. Monogr. 88, 188–203.

Wouters, N., Dakos, V., Edwards, M., Serafim, M.P., Valayer, P.J., Cabral, H.N., 2015. Evidencing a regime shift in the North Sea using early-warning signals as indicators of critical transitions. Estuar Coast Shelf S 152, 65–72.