Title: Evaluating early-warning indicators of critical transitions in natural aquatic ecosystems


Document type: Postprint

Terms of Use: Copyright applies. A non-exclusive, non-transferable and limited right to use is granted. This document is intended solely for personal, non-commercial use.

Evaluating early-warning indicators of critical transitions in natural aquatic ecosystems

Short title: Early-warning indicators in empirical time series

Alena S. Gsell1,2*, Ulrike Scharfenberger1,3, Deniz Özkundakci4, Annika Walters5, Lars-Anders Hansson6, Annette B. G. Janssen2,7, Peeter Nõges8, Philip C. Reid9,10,11, Daniel E. Schindler12, Ellen van Donk2, Vasilis Dakos13 and Rita Adrian1,3

1 Ecosystem Research, Leibniz-Institute of Freshwater Ecology and Inland Fisheries (IGB), Müggelseedamm 310, 12587 Berlin, Germany
2 Aquatic Ecology, Netherlands Institute of Ecology (NIOO-KNAW), Droevendaalsesteeg 10, 6708 PB Wageningen, The Netherlands
3 Freie Universität Berlin, Department of Biology, Chemistry, Pharmacy; Takustr. 3, 14195 Berlin, Germany
4 Waikato Regional Council, 401 Grey St, 3216 Hamilton, New Zealand
5 U.S. Geological Survey, Wyoming Cooperative Fish and Wildlife Research Unit, Department 3166, 1000 E. University Avenue, University of Wyoming, Laramie, Wyoming 82071 USA
6 Institute of Ecology, Lund University,223 62 Lund, Sweden
7 Department of Aquatic Ecology and Water Quality Management, Wageningen University, Wageningen, The Netherlands
8 Centre for Limnology, Institute of Agricultural and Environmental Sciences, Estonian University of Life Sciences, Rannu, 61117 Tartumaa, Estonia
9 Sir Alister Hardy Foundation for Ocean Science, The Laboratory, Citadel Hill, Plymouth PL1 2PB, United Kingdom
10 Marine Institute, Plymouth University, Drake Circus, Plymouth PL4 8AA, UK
11 Marine Biological Association of the UK, The Laboratory, Citadel Hill, Plymouth PL1 2PB, UK
12 School of Aquatic and Fishery Sciences, University of Washington, Box 355020, Seattle, Washington 98195 USA
13 Institute of Integrative Biology, Center for Adaptation to a Changing Environment, ETH Zürich, Universitätstrasse 16, 8092 Zürich Switzerland

* Corresponding author:
Aquatic Ecology, Netherlands Institute of Ecology (NIOO-KNAW), Droevendaalsesteeg 10, 6708 PB Wageningen, The Netherlands Phone: +31 (0) 317 473 553, Email: a.gsell@nioo.knaw.nl

Keywords: alternative stable states, competition, intra-guild predation, resilience indicators, time series, trophic cascade
Abstract

Ecosystems can show sudden and persistent changes in state despite only incremental changes in drivers. Such critical transitions are difficult to predict as the state of the system often shows little change prior to the transition. Early-warning indicators are hypothesised to signal the loss of system resilience and have been shown to precede critical transitions in theoretical models, paleo-climate time series, and in laboratory as well as whole lake experiments. However, the generality of early-warning indicators for detection of critical transitions in empirical time series of natural aquatic ecosystems remains largely untested. Here, we assessed four commonly used early-warning indicators on long-term datasets of five freshwater ecosystems that have experienced sudden, persistent transitions and for which the relevant ecological mechanisms and drivers are well-understood. These case-studies were categorised by three mechanisms that can generate critical transitions between alternative states: competition, trophic cascade, and intra-guild predation. While early-warning indicators could be detected in most case-studies, agreement among the four indicators was low. In some cases, early-warning indicators were detected considerably ahead of the transition. Our results, however, show that, at present, early-warning indicators do not provide reliable and consistent signals of impending critical transitions despite using some of the best routinely monitored freshwater ecosystems. Our analysis strongly suggests that a priori knowledge of the underlying processes driving ecosystem transitions is necessary to identify relevant state variables to successfully monitor early-warning indicators.

Significance Statement

Early-warning indicators are statistical metrics of system resilience and have been hypothesized to provide advance warning of sudden shifts in ecosystems, so-called regime shifts. Here we test this hypothesis for four commonly used early-warning indicators. We used empirical time series from five freshwater ecosystems with documented sudden, persistent transitions hypothesised to represent critical transitions. Early-warning indicators were detected in several of these long-term records, and in some cases, indicators were detected several years before the transition. However, the indicators varied in their reliability, and agreement between indicators was low. Moreover, their applicability was strongly limited by the requirement of ecosystem-specific knowledge of transition generating mechanisms and their drivers to choose relevant state variables for analysis.
Introduction

Ecosystems can show multi-state stability and occasionally sudden transitions from one regime to another despite only incremental changes in drivers (1-5). These critical transitions are characterised by the occurrence of alternative regimes under the same environmental conditions and by abrupt, discontinuous transitions between regimes when a critical threshold is exceeded. As each regime is stabilised by feedback loops, the thresholds for the forward and backward shifts may differ, resulting in hysteresis (5). A well-known example is the nutrient-driven shift between the clear, macrophyte-dominated and the turbid, phytoplankton-dominated regime in some shallow lakes (6). At ecosystem level, such fundamental reorganisations affect ecological processes and hence ecosystem services potentially incurring large economic costs. Therefore, reliable tools to assess ecosystem resilience are sought, ideally providing management with time to avert an impending critical transition (7). Stabilisation through feedback loops, however, often precludes a systematic response in state variables ahead of the shift, rendering prediction of critical transitions difficult (8). The development of early-warning indicators (EWIs) derived from bifurcation theory has drawn considerable interest for detecting critical transitions, particularly for their promise of generality. These EWIs have been shown to precede critical transitions in modelled (8-10), experimental time series (11-13), reconstructed paleo-climate records (14, 15) and whole-lake experiments (16). Up to now, however, an assessment of the generality and the detection power of EWIs on long-term monitoring data in aquatic systems is lacking (17).

Early-warning indicators are statistical metrics that quantify the loss of temporal or spatial resilience and thereby provide advance warning of the potential proximity to a critical threshold (18). Several of these EWIs are related to critical slowing down, a characteristic property of dynamic systems close to catastrophic local bifurcations (19). A bifurcation marks a threshold where the stability properties of the state of the system change. As the system approaches such a threshold, the return rate to equilibrium after a small perturbation slows down so that the system tends to become more similar to its own past, resulting in an increase in autocorrelation at lag-1 (AR-1) (20). This lack of decay of the impact of past perturbations also leads to a build-up in variance, typically measured as an increasing trend in standard deviation (SD) (21). A concurrent increase in SD and AR-1 in a time series also produces higher variability in low frequency processes compared to high frequency processes in the power spectrum of a time series, which
can be quantified as an increasing density ratio (DR) of variance at low to high frequencies (7,22).

In addition to critical slowing down, the system tends to remain for longer at the basin boundary between the two alternative attractors resulting in a skewed distribution of the state variable (SK) (23). However, the expected trends in AR-1, SD, DR, and SK are not exclusive indicators of critical transitions as false positives and false negatives can occur (24). Hence, prior to the application of EWIs based on critical slowing down a careful analysis is needed of whether a system is actually undergoing a critical transition (25).

Abrupt changes in the state of an ecosystem can develop from several mechanisms, including: i) linear tracking of large changes in environmental conditions, ii) non-linear but continuous (reversible) responses to gradual changes in environmental conditions, or iii) non-linear discontinuous (irreversible) responses to gradual changes in environmental conditions (26). While the first mechanism is distinguished by concurrent large changes in environmental drivers (e.g. a sudden increase in temperature), the difference between the other two mechanisms only becomes apparent when the driver is reversed. Hence, single step changes in time series cannot provide direct evidence of a critical transition (27). As empirical time series covering forward and backward shifts are rare, identification of critical transition in natural systems is difficult. However, by linking observed step changes to a mechanistic understanding of the driving processes that can give rise to bi-stability, one can hypothesise (but not prove) the existence of a potential critical transition (5, 28).

Lakes have been proposed as particularly suitable ecosystems to test for EWIs associated to critical transitions. The modular nature of lakes additionally allows comparison across different lakes (29). In aquatic systems, a number of ecological mechanisms have been shown to generate critical transitions between alternative states. The most commonly identified mechanisms include i) competition between two or more species (2, 28), ii) trophic cascades through inclusion or exclusion of top predators (16) or parasites (30) resulting in overexploitation traps, and iii) intra-guild predation through resource competitors that also prey on each other (31, 32). For the purpose of this paper, we selected 14 state variables of five well-documented freshwater case-studies of critical transitions to test whether four commonly used EWIs (AR1, SD, SK and DR) can be detected reliably in advance of the transition. We assessed how often these EWIs showed
the same trends (agreement) and whether their behaviour depended on the mechanism, the type of state variable, the magnitude of the step, and the sampling frequency. Lastly, we also tested how many years ahead of the transition EWIs were detectable. To our knowledge, this study provides the first comprehensive assessment of EWIs in some of the best-documented aquatic time series collected following standard monitoring schemes.

Results

Selection of case-study ecosystems and state variables

We selected five case-study ecosystems based on expert knowledge of well-described regime shifts that can be qualified as critical transitions in aquatic ecosystems: Lake Müggelsee (Germany, LMS), Lake Veluwemeer (The Netherlands, LVM) and Lake Zwemlust (The Netherlands, LZL), Lake Washington (United States, LW), and Lake Võrtsjärv (Estonia, LV). Based on the literature on these case-study ecosystems, we identified relevant critical-transition generating mechanisms, which guided our choice of 14 state variables for EWI analysis. A short summary of the case-studies, mechanisms, state variables and drivers is presented in Table 1. A more detailed description of each case-study and reasoning for the choice of the 14 state variables is presented in the Supplementary Information Appendix (SI Case-studies S1).

Transition detection and seasonal adjustment

We used three complementary methods (piece-wise linear regression, Pettit and STARs, see Methods) to robustly assess the timing of transitions (breakpoints; i.e. large, persistent step changes) in the time series of each state variable. In 12 state variables we found one breakpoint and in two state variables two breakpoints (Table 2 and Figure 1), resulting in 16 pre-breakpoint time series (i.e. from start to step change) for further analysis. In all state variables, the timing of the breakpoints confirmed the timing of transitions reported in the literature. Information on the data structure including time period, sampling interval, number of data points and the percentage of missing values in the pre-breakpoint time series are summarised in SI Table S2. Each state variable showed significant differences in the pre- to post-breakpoint means (Table 2), with step magnitudes ranging from 0.24*standard deviation to 1.59*standard deviation (Table 2). Each pre-breakpoint time series was detrended and seasonally adjusted using a Gaussian smoother with a bandwidth corresponding to 12 (for monthly data sets) or 26 (for fortnightly datasets) data points.
Testing the residual time series for remaining linear trends and seasonality showed in some time series a remaining, but much reduced seasonal signal (SI Table S2).

Early-warning indicator analysis

In 14 out of 16 analysed time series, a loss of resilience before the breakpoint was signalled by at least one of the four EWI metrics (coded ‘+’ in Table 2, detailed figures in SI Figures S3) based on the median of the trend distribution across yearly increments of rolling window sizes (see Methods). We found rising AR-1 and DR trends in 10 cases each (63 %), rising SD trends in 9 cases (56 %), and increasing or decreasing SK trends (according to the direction of the state change) in 7 cases (44 %) (Figure 2, Table 2). In some cases, AR-1 trends were increasing from negative values to positive ones (see SI Figures S3). We counted these trends as positive (coded ‘(+)’ in Table 2). In several cases, indicators showed trends opposite to the theoretical expectation (coded ‘–’ in Table 2). Only the state variable phytoplankton biomass in Lake Müggelsee (LMS) showed the theoretically expected trends in all EWIs, while in two state variables (non-algal attenuation in Lake Veluwemeer (LVM) and cyanobacteria biomass in Lake Washington (LW)) all EWIs failed. The agreement between positive AR-1 and SD trends was low (5 cases), but higher between positive AR-1 and DR trends (10 cases). Logistic regressions showed no significant relationship ($p < 0.05$) between EWI behaviour and mechanism (competition, IGP, or trophic cascade), state variable level (species, group, or ecosystem), step change height, length of pre-breakpoint time series, nor sampling interval (fortnightly or monthly).

Robustness to rolling window size and significance of EWI trends

We estimated the robustness of EWI trends to the size of the rolling window, shown as the distribution of trends around their median (boxplots in Figure 2 for AR-1 and SD and SI Figure S4 for DR and SK). A large majority of trends were robust to rolling window size, as shown in the low number of trend distributions in which the boxplot extended to negative values. We also tested for the significance of the trends by estimating the rate of false positives using simulated surrogate time series (see Methods). Out of all 64 EWI trends (16 time series by four EWIs) only three trends were significant in more than 50% of the comparisons between data-based and surrogate-based trends. Additional 29 EWI were significant in less than 50% of comparisons, and 32 EWI showed no significant differences between data-based and surrogate-based trends (grey
Early warning indicators in stepwise shortened time series

To assess whether EWI trends could have been detected in incomplete time series that would have been available one or several years before the actual transition took place, we repeated the EWI analysis on stepwise shortened (yearly increments) time series of each state variable. In 14 (out of 16) time series, at least one of the EWIs would have indicated a loss in ecosystem resilience 1 to >10 years before the critical transition. In three instances positive EWIs trends occurred just the year before the shift (SI Table S5), while in 17 instances the median of the EWI trend distribution remained positive (i.e. showing the expected sign) up to the minimum time series period that was analysed (three years).

Discussion

In this study we assessed the detectability of four commonly used EWIs (AR-1, SD, SK, and DR) in empirical time series of freshwater ecosystems. By combining high quality empirical time series with ecological understanding and standardised methods we showed that EWIs preceded critical transitions in natural aquatic ecosystems, in some cases even several years ahead of the shift, despite potential shortcomings of empirical datasets such as observation error, sparse sampling or low signal to noise ratio (9, 24, 33). However, in a large proportion of cases the EWIs failed, and the cases with positive EWIs generally showed low or no significance. In addition, we found little agreement among signals and we observed no relationship between EWI trends and potential predictors, like ecological mechanism of shift, state variable level, magnitude of transition, or sampling interval.

Use of ecological understanding in the choice of state variables

The choice of case-studies and state variables in our study was based on expert knowledge of ecosystems that likely experienced critical transitions. Although this choice was based on ecological understanding of the mechanisms that can give rise to alternative states in aquatic ecosystems, it does not provide conclusive evidence that the regime shifts we analysed correspond to true critical transitions. Still, such a priori choices of state variables for EWI analysis implicitly excluded assumptions based on other transition types, such as responses to step changes in the
driver (34), but offered clear expectations of what EWI behaviour should precede a transition (25). As mechanisms generating critical transitions can operate at all ecosystem levels, relevant state variables range from discrete variables, such as species-specific biomass (e.g. (32)), to aggregated variables, such as Secchi depth, turbidity or metabolism (e.g. (12)). Additionally, the case-studies were chosen based on the availability of long time series of monitoring data of relevant state variables at sufficiently high temporal resolution and with few missing values (35). Despite our conscientious choice of ecosystems and state variables, we still found contradicting patterns in EWIs.

**Detection and agreement among indicators**

Not all EWIs were equally reliable in detecting impending transitions (Table 2). In 44 % of the total 64 cases the EWI failed. Only in 3 cases where signals were positive, trends were significant in more than 50% of data-based and surrogate-based trend comparisons. Our significance testing relied on estimating EWI trends in surrogate stationary data fitted to the original time series to determine the rate of false positives. Unfortunately, we lack records to act as controls for comparing trends from comparable aquatic systems where no transition took place to estimate the rate of true negatives (no alarm, or sensitivity) of the EWIs. Given the lack of such controls, one potential way to measure the no alarm rate is to derive trends from non-stationary models fitted to the data (36). It would be valuable in future work to compare trends in EWI in study systems that either did not show any changes or in which changes were brought about by large external shifts in drivers.

Agreement between AR-1 and SD trends has been postulated as a minimum requirement to signal the approach of a transition (8). In our dataset, these two indicators concurrently increased in only 5 out of 16 time series. Critical transitions with increasing AR-1 and decreasing SD have also been observed in other studies (36). Such inconsistent trends between indicators may occur in ecosystems that are subject to multiple concurrent regime-shift generating processes that may or may not interact via shared state variables and that may react differently to drivers and environmental noise (9). If such connected regime-shift processes work towards muffling variance in the measured state variable, the variance-based EWI signal may be suppressed (37). Transitions from cycles to stable points can also generate decreasing variation (38) and may explain
decreasing SD trends like in the Lake Müggelsee *Cyclops vicinus* abundances or the Lake Võrtsjärv functional group U biomass.

In many cases autocorrelation (AR-1) was generally low and in some cases even negative (e.g. Lake Washington non-*Daphnia* cladocerans). Low AR-1 values (far from the theoretical value of 1 where critical transitions occur) reflect that transitions in the real world are likely triggered well before the actual tipping point is reached (25, 39). Evaluating the full power spectrum for changes in power in aggregated low versus high frequencies over time indicated that higher order AR processes did not provide more information than the AR-1. This suggests that the reported low or negative autocorrelation in our records probably originated from the too long sampling intervals in cyclic variables (e.g. population cycles) resulting in under-sampled cycles of fast growing plankton and its related variables.

Successfully detecting EWIs has often been related to the availability of high sampling frequency data (40), although it has been shown that EWIs could still be detected robustly in infrequently sampled data as long as the time series were sufficiently long (41). Our study was based on data sampled at (or averaged to) fortnightly and monthly intervals. Monthly intervals are rather long compared to the generation or reaction times of the state variables tested in our study (e.g. phytoplankton generation times are in the order of days, therefore monthly averages of biomass aggregate multiple generations of phytoplankton). Nevertheless, we could not detect a relationship between sampling interval and the proportion of failing EWIs in our dataset. Finally, data preparation, in particular choices on detrending and seasonal adjustment methods, can affect the autocorrelation structure in time series and hence the outcome of EWI analyses. We used Gaussian filtering with a fixed bandwidth based on the sampling frequency to detrend and seasonally adjust the data before EWI analysis (see SI S6 for a comparison of seasonal adjustment methods). Despite data preparation, some time series still showed reduced traces of seasonality. Remaining seasonal signals may increase or decrease the intercept of the EWI trends but not the sign of the trend.

*Early detection based on incomplete time series*

Informative changes in some of the EWI metrics were already detectable several years preceding
the actual transition, although large differences in detection windows between state variables were observed. These differences may partly be explained by the fact that the actual shift is usually triggered by external stochastic perturbations which in turn are often independent of the drivers of ecosystem stability loss (42). However, in many cases EWIs indicated sustained instability over the period tested in our study which may be attributed to either a too short time span available for testing (e.g. Lake Zwemlust) or unrecognised interacting processes that promoted prolonged instability in these ecosystems.

Conclusion
Despite our informed and conscientious choice of case-study ecosystems and state variables, we found relatively low detectability of EWIs prior to the documented transitions, and when EWIs were detected, the agreement among EWIs was low. Our findings are in line with results from an assessment of the detectability of EWIs prior to non-linear transitions (43). Although it is encouraging that we could detect EWIs in some of our empirical aquatic time series using data derived from commonly used monitoring schemes that were not designed for this purpose, the lack of reliability and agreement between signals limits the potential application of EWIs to well-understood ecosystems only (35). In such well-understood ecosystems, harnessing EWIs as metrics of resilience loss may help in planning for the unpredictable and could be part of strategic foresight programs for management and conservation (44). However, our analysis suggests that these metrics could be of added value only in combination with existing frameworks (e.g. alternative stable state theory) and in-depth ecosystem knowledge. Furthermore, taking into account the underlying assumptions and requirements of EWI analysis can inform managers about adaptations in monitoring schemes by advising about relevant variables and temporal sampling resolution to adequately capture changes in the resilience of systems. One way forward may be the advent of automated, continuous high-frequency monitoring (35), ideally monitoring multiple lakes with similar properties for comparison (35). To increase our understanding of critical transition generating processes, ecosystem models, such as PCLake, can help to bridge the gap between simple minimal models and the full complexity of natural systems and allow combined analysis of e.g. food-web theory and alternative stable states theory and indicators of ecosystem resilience (45). Ideally, such insights can serve to broaden our search image in empirical EWI patterns instead of relying on a generic increase in variance without understanding the inherent
variability in ecosystems. In the meantime, the reliability of EWIs for predicting abrupt shifts in ecosystem state should be treated with caution.

Methods

Data preparation and breakpoint detection

The data preparation and statistical analysis protocol was identical for all 16 state variable time series. Each time series was analysed at the highest temporal resolution available or at a lower resolution that resulted in fewer missing data (see SI Table S2). Lakes Müggelsee and Washington were analysed at fortnightly, all other case-studies at monthly intervals. As our time series methods require continuous and equidistant data, we imputed missing values up to a maximum of four consecutive time steps using a Kalman filter (46). Time series with longer gaps were shortened to start or end at the gap (e.g. Lake Võrtsjärv). Each time series was standardized by mean centering and standard deviation scaling for convenient comparison of step change heights.

The timing of the step was determined by breakpoint analysis as step changes in the respective state variables may differ from the timing of whole-system step changes reported in the literature. Robust estimates of the timing of step change were achieved by employing three complementary breakpoint estimation methods: a) additive decomposition of time series in seasonal, trend and residual components and subsequent iterative fitting of piecewise linear season and trend models (47) using the R package “bfast” (48), b) testing for step changes in the average using the Pettitt test (49) and c) STARS, a combination of a sequential partial CUSUM method and a t-test (50). If at least two methods showed similar timing for a step change (± 12 months), the standardized original time series was split at that breakpoint. Differences between pre- and post-breakpoint means of the time series were tested with a Welch two sample t-test (51) (Table 2).

Due to a one-year gap in the time series, breakpoint timing of two state variables (LV functional groups P (eutrophic epilimnion species) and U (summer epilimnion species)) was determined differently: here we assessed whether large changes in the biomasses of the functional groups occurred during the gap indicating a potential shift. The pre-1978 and post-1978 time series showed significantly different means for functional group U (Welch two-sample t-test, t=3.52, df=128, p< 0.001, difference in means = 0.29*standard deviation) and functional group P (Welch two-sample t-test, t=-4.67, df=439, p< 0.001, difference in means = 0.55*standard deviation)
suggesting that the shift in functional groups U and P occurred during the year 1978 (Table 2). As both of these time series showed no further breakpoints in the years after 1978, we conservatively assumed 1977 to be the transition year.

Early-warning indicator analysis

The pre-breakpoint time series was seasonally adjusted using a Gaussian smoother with a kernel bandwidth based on the number of data points per year (i.e. 12 for monthly and 26 for fortnightly datasets, see SI S6 for comparison of other methods). A bandwidth of one year was chosen to account for yearly recurring patterns while retaining informative low and high frequency variability other than long-term trend and season (15). In three residuals time series, local outliers were replaced by Kalman imputed values (LMS Dreissena (two outliers) and Aphanizomenon Period 1 and LW Cyanobacteria (one outlier each)). The residuals were then passed on to analysis of EWIs AR-1, SD, SK, and DR with testing for robustness to the size of rolling window and testing for significance (false positives) using the R package “earlywarnings”(9). Additionally, we confirmed that the static choice of fixed compared frequencies in the EWI density ratio (DR) was sufficiently capturing changes of the full power spectrum of the pre-breakpoint residuals time (based on its estimated smoothed Fast Fourier Periodogram).

Robustness and significance testing

The trends in the estimated temporal evolution of EWIs from the rolling-window approach were quantified by the non-parametric Mann-Kendall trend test, which tests for monotonic trends based on the Kendall τ rank correlation coefficient (9). As the size of the rolling window can affect the EWI trends (36), a robustness analysis was performed estimating the distribution of trends and proportion of trends that did not differ in sign from the median of the trend distribution. We did this by using yearly increments of the residuals time series covered by the rolling time-window (two to n-two years, function “sensitivity_ews” (9), where n was the number of years in each pre-breakpoint time series). Significance testing was conducted by comparing the data-based EWI trend against a bootstrapped distribution of 200 surrogate time series-based EWI trends. The surrogate time series were generated based on an ARMA(p,q) model fitted on the residuals time series. The bootstrapped distribution of trends depicts the probability that a particular trend could occur by chance in time series of the same ARMA structure (function “surrogates_ews” in R-
package “earlywarnings” (9)). The data-based EWI trend was deemed significant if it fell on one of the 5% tails of the surrogate-based trend distribution (α=0.1). This significance testing was repeated for all rolling window sizes and the proportion (%) of significant trends over all rolling window sizes was reported.

**Relationship of EWIs with state variable categories and agreement among signals**

The relationship between the occurrence of each EWI and predictor variables mechanism (competition, IGP or trophic cascade), state variable category (species, group, ecosystem), step change height, length of pre-breakpoint time series and sampling interval (fortnightly or monthly) was tested with logistic regressions (identity link for continuous predictors (step change height and time series length); logit link for categorical predictors) with Bonferroni corrected post hoc testing. The median of the trend distribution across all time-window sizes was used to assess agreement among EWIs.

**Early detection of early-warning indicators**

To assess how timely EWIs could have detected the approaching transition, we quantified EWI trends on stepwise shortened pre-breakpoint time series, starting with the full time period and continuing with stepwise reduction of the time series by cutting off the last year of data. The minimum time series length was set to three years to allow meaningful Kendall τ estimation. How many years before the shift an EWI could already have been detected was set by assessing the maximum number of stepwise reductions before the theoretically expected EWI trend disappeared (i.e. when the median of the Kendall τ trend distribution turned negative for AR-1, SD and DR or changed sign for SK). All data analyses and graphing were conducted using the R language environment for statistical computing (52) and associated library extensions.

**Acknowledgements**

ASG, RA and LAH were supported by the EU-project LIMNOTIP funded under the FP7 ERA-Net Scheme (Biodiversa, 01LC1207A) and LakeShift (DFG). ABGJ was funded by the Netherlands Organization for Scientific Research (NWO) project no. 842.00.009. VD acknowledges a Rubicon fellowship from the Netherlands Science Foundation (NWO), a EU Marie Curie grant, and an ACE fellowship by ETH Zurich. The authors would like to
acknowledge Rijkswaterstaat as source for the Lake Veluwemeer dataset. We thank Bryan Spears, Francis Daunt, Sarah Burthe, Silke Schmidt, Erik Jeppesen, Brian Moss and the members of LIMNOTIP for valuable discussions. The authors also thank the anonymous reviewers and the subject editor for their constructive and positive comments which helped to improve this manuscript considerably. We also thank the Federal Ministry of Education and Research (BMBF) for their administrative support. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.
Figure captions

Figure 1: Time series of 14 selected state variables (black lines: competition, green lines: trophic cascade, blue line: intra-guild predation) of five lakes. Solid red lines indicate the timing of critical transitions (estimated by breakpoint analysis, see Methods). Dashed red lines indicate additional data-based breakpoints that were not used in this study as they were either not described in the literature (LMS phytoplankton biomass) or resulted from bio-manipulation (LZL Secchi depth).

Figure 2: Robustness and Significance testing. Mann-Kendall trend distributions for EWIs AR-1 and SD (panel) and per analysed time series (boxplots) and the proportion of significant differences in trends between data-based and surrogate time series across all rolling window sizes (significance, grey bar plots). Positive values of trends in the boxplots represent increasing EWI trends. Note almost in all time series the trends were robust to the choice of rolling window size. Significance testing was based on comparing indicator trends produced from stationary surrogate time series to the empirically reported for fixed rolling window size (see Methods).

Table captions

Table 1: Summary of case-study systems, state variables and drivers by mechanism: trophic cascade, intra-guild predation and competition

Table 2: Overview for all 14 state variables (16 time series) listing the direction of the shift, timing of the break-point (year-month), step height (scaled to standard deviation), the two-sample Welch test significance, and for each EWI the sign of the median of the trend distribution (across all rolling window-sizes). Bold t (Welch) values are significant at $\alpha=0.05$ level. The tested EWI are autocorrelation at lag-1 (AR-1), standard deviation (SD), skewness (SK), and density ratio (DR). Median trends that corresponded with the theoretical expectation were coded with ‘+’, trends that showed the opposite to expectation with ‘-’. Brackets denote AR-1 trends crossing zero on the scale.
References


51. Welch BL (1947) The generalization of student's problem when several different population variances are involved. *Biometrika* (34) 28-35.


Table 1: Summary of case-study systems, state variables and drivers by mechanism: trophic cascade, intraguild predation and competition

<table>
<thead>
<tr>
<th>Mechanism</th>
<th>Case-Study</th>
<th>Shift in State Variable</th>
<th>Driver</th>
<th>Process</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trophic cascade</td>
<td>LW</td>
<td>Increase in water transparency</td>
<td>Increase in grazing pressure</td>
<td>Longfin smelt (<em>Spirinchus thaleichthys</em>) preys on secondary consumer <em>Neomysis</em> which releases primary consumer <em>Daphnia</em> who grazes on producers (phytoplankton)</td>
<td>(53, 54)</td>
</tr>
<tr>
<td></td>
<td>LMS</td>
<td>Increase in <em>Dreissena polymorpha</em> larvae</td>
<td>Decrease in predation pressure</td>
<td>Changes in dominant carnivorous zooplankton species coincided with increase in <em>Dreissena</em> larvae.</td>
<td>(55, 56)</td>
</tr>
<tr>
<td></td>
<td>LMS</td>
<td>Decrease (period 1) and increase (period 2) in <em>Leptodora kindii</em></td>
<td>Increase (period 1) and decrease (period 2) in predation pressure</td>
<td><em>Leptodora</em> are a preferred prey for fish and likely indicate changes in overall fish predation pressure</td>
<td>(56)</td>
</tr>
<tr>
<td>Competition</td>
<td>LMS</td>
<td>Decrease phytoplankton biomass</td>
<td>Re-oligotrophication</td>
<td>Reduction in nutrients decreases phytoplankton growth, improving light climate favouring macrophyte reestablishment</td>
<td>(57, 58)</td>
</tr>
<tr>
<td></td>
<td>LMS</td>
<td>Decrease in <em>Aphanizomenon</em> (period 1)</td>
<td>Re-oligotrophication counteracted by spring warming</td>
<td>Warmer springs promote cold-adapted cyanobacteria development</td>
<td>(59)</td>
</tr>
<tr>
<td></td>
<td>LMS</td>
<td>Decrease in <em>Aphanizomenon</em> (period 2)</td>
<td>Re-oligotrophication</td>
<td>Reduction in nutrients decreases phytoplankton growth, improving light climate favouring macrophyte reestablishment</td>
<td>(57)</td>
</tr>
<tr>
<td></td>
<td>LW</td>
<td>Decrease in non-<em>Daphnia</em> cladocera</td>
<td>Reduction in predation pressure on <em>Daphnia</em></td>
<td>Indirect effect of trophic cascade through increasing resource competition by <em>Daphnia</em></td>
<td>(53, 54)</td>
</tr>
<tr>
<td></td>
<td>LW</td>
<td>Increase in cryptophytes and decrease in cyanobacteria</td>
<td>Re-oligotrophication</td>
<td>Reduction in nutrients decreases cyanobacteria competitive ability and releases other phytoplankton from competition</td>
<td>(53, 54)</td>
</tr>
<tr>
<td></td>
<td>LZL</td>
<td>Decrease in water transparency</td>
<td>Eutrophication, epiphyte shading and herbivory reduction in nutrients and bentivorous fish</td>
<td>Competition between submerged vegetation and phytoplankton under eutrophication and herbivory on macrophytes</td>
<td>(60, 61)</td>
</tr>
<tr>
<td></td>
<td>LVM</td>
<td>Decrease in non-algal attenuation</td>
<td>Eutrophication</td>
<td>Recovery of submerged vegetation cover and subsequent stabilisation of sediments</td>
<td>(62)</td>
</tr>
<tr>
<td></td>
<td>LV</td>
<td>Increase in functional group P</td>
<td>Eutrophication</td>
<td>Competition between functional groups P (eutrophic epilimnion species) and U (summer epilimnion species)</td>
<td>(63, 64)</td>
</tr>
<tr>
<td></td>
<td>LV</td>
<td>Decrease in functional group U</td>
<td>Eutrophication</td>
<td>Competition between non-nitrogen fixing and di-nitrogen fixing species (functional group H1)</td>
<td>(63, 64)</td>
</tr>
<tr>
<td>Intraguild predation</td>
<td>LMS</td>
<td>C. vicinus - C. kolensis dominance switch</td>
<td>Reduction of shared food source</td>
<td>The inferior resource competitor <em>C. vicinus</em> preys on juveniles of the smaller <em>C. kolensis</em></td>
<td>(32)</td>
</tr>
</tbody>
</table>
Table 2: Overview for all 14 state variables (16 time series) listing the direction of the shift, timing of the transition (year-month), step height (scaled to standard deviation), the two-sample Welch test significance, and for each EWI the sign of the median of the trend distribution (across all rolling window-sizes). Bold t (Welch) values are significant at $\alpha=0.05$ level. The tested EWI are autocorrelation at lag-1 (AR-1), standard deviation (SD), skewness (SK), and density ratio (DR). Median trends that corresponded with the theoretical expectation were coded with ‘+’, trends that showed the opposite to expectation with ‘-’. Brackets denote AR-1 trends crossing zero on the scale.

<table>
<thead>
<tr>
<th>Case-Study</th>
<th>State Variable</th>
<th>Direction of Shift</th>
<th>Break-Point (Year-Month)</th>
<th>Step (scaled to SD)</th>
<th>t (Welch)</th>
<th>AR-1</th>
<th>SD</th>
<th>SK</th>
<th>DR</th>
</tr>
</thead>
<tbody>
<tr>
<td>LMS</td>
<td>Phytoplankton mg L$^{-1}$</td>
<td>Decrease</td>
<td>1990-5</td>
<td>0.803</td>
<td>9.92</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td><em>Aphanizomenon</em> mg L$^{-1}$ P1</td>
<td>Decrease</td>
<td>1990-11</td>
<td>0.413</td>
<td>4.47</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td><em>Aphanizomenon</em> mg L$^{-1}$ P2</td>
<td>Decrease</td>
<td>2002-12</td>
<td>0.317</td>
<td>5.92</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td><em>Cyclops vicinus</em> ind L$^{-1}$</td>
<td>Decrease</td>
<td>1992-6</td>
<td>0.629</td>
<td>7.86</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td><em>Dreissena polymorpha</em> larvae ind L$^{-1}$</td>
<td>Increase</td>
<td>1993-9</td>
<td>0.388</td>
<td>-6.79</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td><em>Leptodora kindtii</em> ind L$^{-1}$ P1</td>
<td>Decrease</td>
<td>1987-11</td>
<td>0.776</td>
<td>6.58</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td><em>Leptodora kindtii</em> ind L$^{-1}$ P2</td>
<td>Increase</td>
<td>2005-7</td>
<td>0.239</td>
<td>-5.21</td>
<td>(+)</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>LW</td>
<td>Secchi depth m</td>
<td>Increase</td>
<td>1976-10</td>
<td>1.29</td>
<td>-21.94</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Cryptophyceae 100 µm$^{-1}$ L$^{-1}$</td>
<td>Increase</td>
<td>1977-7</td>
<td>0.627</td>
<td>-13.16</td>
<td>(+)</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Cyanophyceae 100 µm$^{-1}$ L$^{-1}$</td>
<td>Decrease</td>
<td>1973-4</td>
<td>1.587</td>
<td>9.33</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Non-<em>Daphnia</em> cladocerans ind L$^{-1}$</td>
<td>Decrease</td>
<td>1976-8</td>
<td>0.783</td>
<td>6.09</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>LV</td>
<td>H1 g/m$^{3}$</td>
<td>Increase</td>
<td>1992-8</td>
<td>0.569</td>
<td>-6.75</td>
<td>(+)</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>P g/m$^{3}$</td>
<td>Increase</td>
<td>1977-12</td>
<td>0.281</td>
<td>-4.61</td>
<td>(+)</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>U g/m$^{3}$</td>
<td>Decrease</td>
<td>1977-12</td>
<td>0.532</td>
<td>3.17</td>
<td>(+)</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>LVM</td>
<td>Non algal attenuation m$^{-1}$</td>
<td>Decrease</td>
<td>1995-9</td>
<td>1.058</td>
<td>9.05</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>LZL</td>
<td>Secchi depth m</td>
<td>Decrease</td>
<td>1992-9</td>
<td>1.098</td>
<td>7.97</td>
<td>(-)</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

Proportion of EWI trend distribution medians corresponding to the theoretical trend expectation

<table>
<thead>
<tr>
<th></th>
<th>10/16</th>
<th>9/16</th>
<th>7/16</th>
<th>10/16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 1

[Graphs and data plots related to phytoplankton, Aphanizomenon, C. vicinus, D. polymorpha, L. kindti, LVM attenuation, and Secchi depth over years from 1970 to 2010 for different locations and types of data.]
Figure 2

[Boxplot diagram showing AR-1, SD, and % sig for various categories including LMS, LW, LV, LVM, LZL. Categories include Phytoplankton, Aphanizomenon P1, A. winogradovii, D. polymorpha, L. kindtii P1, Secchi, Cryptophytes, Cyanobacteria, ND cladocerans, Functional group H1, Functional group P, Functional group U, Attenuation, and Secchi.]