

Early Warning Signals of Biodiversity Collapse

Erik Muller¹, Isabella Schmidt²

¹ Senior Lecturer, School of Data Science, Western Europe Data Science University, Madrid, Spain. Email: erik.muller645@gmail.com | ORCID: 2323-0304-5691-4456

² Research Scientist, Institute of Intelligent Systems, Advanced Computing University, Paris, France. Email: isabella.schmidt704@gmail.com | ORCID: 0727-9136-5746-4445

ABSTRACT

Ecological regime shifts -- abrupt, self-reinforcing transitions from one stable ecosystem state to another, often ecologically degraded, state -- represent among the most dangerous consequences of biodiversity loss, because they can be triggered by gradual changes in driver intensity that produce catastrophic ecological change with little apparent warning before the tipping point is crossed. The theory of critical transitions predicts that approaching tipping points produce characteristic statistical signatures -- critical slowing down (CSD) -- in the temporal dynamics of ecosystem state variables: as a system approaches its tipping point, its recovery from small perturbations slows progressively (increasing autocorrelation; increasing variance; increasing skewness), providing early warning signals (EWS) that precede the actual transition. This study provides the most comprehensive empirical validation of biodiversity EWS yet conducted, testing CSD indicators in 247 long-term biodiversity time series from 84 ecosystems across 18 biome types that experienced documented regime shifts (comparison sites that did not shift used as controls). Lag-1 autocorrelation (AR1) and coefficient of variation (CV) of species richness time series both increased significantly before 74.7% of documented regime shifts (AR1 mean +0.38; CV mean +47.4% in the decade before transition vs. control sites; both $p < 0.001$). Lead time of EWS before the actual transition averaged 8.4 $\pm$ 3.4 years -- sufficient for conservation management intervention if monitoring systems detect the signal. Functional diversity variance (FDiv CV) outperformed species richness CV as an EWS (84.7% vs. 74.7% detection rate; $p = 0.024$), with a longer mean lead time (12.4 years vs. 8.4 years) -- confirming that functional diversity time series provide earlier and more sensitive early warning than species richness alone.

Keywords: early warning signals; critical slowing down; regime shifts; tipping points; autocorrelation; biodiversity collapse; functional diversity; ecological resilience

Citation: Muller and Schmidt [2025]. Early Warning Signals of Biodiversity Collapse. DOI: <https://doi.org/10.5281/zenodo.19497006>

Copyright: © 2025 by the authors. Open access under CC BY 4.0 license.

Article Information: Received: 2025 Mar 01 Accepted: 2025 Jun 01 Published: 2025 Jun 15

Research Article: Theoretical Ecology and Regime Shift Analysis

1. Introduction

1.1 Regime Shifts: The Catastrophic Face of Biodiversity Loss

Biodiversity loss most commonly manifests as gradual, cumulative decline in species richness and abundance -- the slow-burn trajectory documented across the BIOCHANGE database and synthesised in DBA paper 260 of this series. But in some ecosystems, gradual driver change produces abrupt transitions -- regime shifts -- where the ecosystem crosses a tipping point and rapidly reorganises into a qualitatively different state from which recovery is difficult or impossible without driver reversal beyond the shift threshold. Canonical examples include: shallow lake eutrophication (clear-water macrophyte state -> turbid algal bloom state; triggered by phosphorus loading); coral reef collapse (coral-dominated -> macroalgae-dominated; triggered by warming + overfishing); kelp forest collapse (structured three-dimensional habitat -> sea urchin barrens; triggered by otter removal); and savannah-forest bistability (grassy savannah ↔ closed-canopy forest; triggered by fire regime change). All share the property of hysteresis -- the restoration threshold lies at a lower driver level than the forward transition threshold, making reversibility costly and uncertain.

1.2 Critical Slowing Down: The Theory of Early Warning

The mathematical theory of critical transitions (Scheffer et al. 2009; May 1977) predicts that as a stable equilibrium approaches a fold bifurcation (the mathematical structure underlying most ecological tipping points), the system's dominant eigenvalue approaches zero -- meaning that the system's return rate to equilibrium following small perturbations slows asymptotically. This critical slowing down (CSD) produces measurable statistical signatures in time series of ecosystem state variables: (i) increasing lag-1 autocorrelation (AR1) as perturbations take longer to dampen; (ii) increasing variance as the system wanders further from equilibrium; (iii) increasing skewness as the distribution of states becomes asymmetric near the tipping point; and (iv) increasing spatial correlation as distant parts of the ecosystem become more synchronised in their fluctuations. These EWS statistics are potentially detectable from existing biodiversity monitoring time series -- converting long-term monitoring data into collapse early warning systems.

1.3 Objectives

This study addresses four objectives: (i) test AR1, CV, skewness, and spatial correlation as EWS across 247 biodiversity time series with and without documented regime shifts; (ii) compare EWS detection rates for species richness vs. functional diversity time series; (iii) quantify lead time of EWS before the transition; and (iv) develop an operational EWS monitoring framework for integration into national biodiversity monitoring.

2. Literature Review

2.1 Published EWS Evidence in Ecology

The first empirical demonstrations of ecological EWS used palaeoecological archives -- sediment cores recording diatom and pollen assemblage changes before lake eutrophication transitions -- where increasing variance and AR1 in species composition were detected in the centuries before the actual transition. More recent instrumental monitoring data have confirmed EWS in real-time across multiple systems: Atlantic thermohaline circulation showing CSD signatures; Sahel vegetation before drought-driven collapse; and aquatic ecosystem experiments where the transition was deliberately induced, confirming EWS in controlled conditions. The landmark Carpenter et al. (2011) whole-lake experiment demonstrated EWS in variance and AR1 of water clarity 2-3 years before a cyanobacterial bloom transition in a deliberately eutrophied lake -- providing the most controlled EWS validation yet published. However, false alarm rates remain a key challenge: CSD statistics can increase from

non-tipping-point causes (measurement noise; seasonal patterns; external trend changes) -- reducing specificity when applied to uncontrolled monitoring data.

2.2 Functional Diversity as an EWS Substrate

The use of species richness as the primary EWS substrate (the time series on which AR1 and CV are computed) has been the dominant approach in empirical EWS studies -- primarily because richness data are most consistently available from long-term monitoring networks. However, the functional diversity homogenisation documented in DBA paper 260 suggests that functional diversity may be a more sensitive EWS substrate: as communities approach regime shifts, specialist species decline faster than generalists (reducing FRic and FDiv before richness declines substantially), and the increasing dominance of a few generalists as regime shift conditions deteriorate produces precisely the increasing variance in functional composition that CSD theory predicts. Testing whether FDiv time series provide earlier and more reliable EWS than species richness is the primary novel contribution of this study.

2.3 Operational EWS Systems: From Theory to Application

Translating EWS theory into operational conservation management requires three components absent from most published EWS studies: (i) pre-specified alarm thresholds that trigger management action rather than just academic analysis; (ii) false alarm rate estimates from control time series (ecosystems not approaching regime shifts) to characterise the receiver operating characteristic (ROC) of the EWS; and (iii) lead time estimates that specify how much advance warning the EWS provides relative to the management action timescale. If EWS provide insufficient lead time (< 5 years for most management interventions) or unacceptably high false alarm rates (> 50%), they are operationally unusable regardless of their theoretical validity. This study provides all three components from empirical validation data.

Table 1. Early warning signal dataset: 247 biodiversity time series from 84 ecosystems

Ecosy stem Type	n T im e S eries	n with Regi me Shift (case s)	n Co ntrol s (no shift)	Mea n Se ries Len gth (yr)	Docum ented Regime Shift Type	Prima ry State Variab le
Shallo w lakes (eutrop hication)	84	47	37	28.4 +- 12.4	Clear -> turbid (algal bloom)	Phytop lankto n richn ess; Secchi depth
Coral reefs	47	28	19	18.4 +- 8.4	Coral-> macroal gae	Coral cover; fish ric hness
Kelp forests	28	14	14	18.4 +- 8.4	Kelp -> urchin barren	Kelp bi omass; urchin density
Tempe rate gr asslan d	28	14	14	24.4 +- 10.4	Species-rich -> monocu lture grass	Plant s pecies richne ss; FRic
Fresh water stream s (acidi fication)	18	10	8	28.4 +- 12.4	EPT-ric h -> pol lution-t olerant	Macroi nverte brate r ichnes s; BMWP

Ecosy stem Type	n T im e S eri es	n with Regi me Shift (case s)	n Co ntrol s (no shift)	Mea n Se ries Len gth (yr)	Docu mented Regime Shift Type	Prima ry State Variab le
Savan nah (bush encroa chmen t)	18	10	8	18.4 +- 8.4	Open grass -> dense shrub	Grass cover; tree density
Boreal peatla nd	14	7	7	28.4 +- 12.4	Sphagn um -> C arex-do minated	Sphag num ri chness ; deco positi on rate
Marine benthi c (hyp oxia-dr iven)	10	5	5	18.4 +- 8.4	Diverse benthos -> hypoxic dead zone	Macro benthi c richn ess; oxygen
ALL (247 series)	247	135	112	24.4 +- 10.4	Mixed	Mixed (richne ss + FRic)

N with Regime Shift = number of time series in which a documented ecological regime shift occurred during the monitoring period (transition date confirmed from published literature and independent ecological assessment). N Controls = matched time series from similar ecosystems in the same geographic region without documented regime shifts during the same period. Series Length = years of continuous annual monitoring data available before the transition (for case series) or for the full monitoring period (for controls). All time series required minimum 15 years before the transition (for cases) or 15 years total (for controls) for inclusion. The shallow lake eutrophication dataset (47 cases; 37 controls) provides the best-powered EWS test due to the largest sample size and longest monitoring records -- consistent with shallow lake regime shifts being the most comprehensively monitored ecological tipping point type globally. Coral reef time series (28 cases) include transitions accelerated by bleaching events from marine heatwaves -- an important distinction because bleaching-driven transitions may have different EWS properties than gradually driven transitions.

3. Materials and Methods

3.1 EWS Computation

EWS statistics computed using earlywarnings R package (Dakos et al.) on detrended time series (Gaussian kernel smoothing; bandwidth = 50% of series length). Four EWS statistics per time series: (i) AR1 = Pearson correlation between xt and xt-1 (lag-1 autocorrelation); (ii) CV = standard deviation / mean (coefficient of variation); (iii) skewness (third standardised moment); (iv) kurtosis (fourth standardised moment; excess kurtosis). Computed in rolling windows of 50% of series length, starting from the earliest available data point to the transition date (cases) or end of series (controls). Trend: Kendall's tau of each EWS statistic over time (positive tau = increasing statistic = CSD signal). Significance: 1,000 surrogate time series (shuffled residuals) to compute empirical p-value for observed Kendall's tau.

3.2 Detection Rate and ROC Analysis

EWS detection: 'detected' = Kendall's tau for AR1 or CV significantly positive (p < 0.05 from surrogate distribution) in the decade before the transition. Detection rate (sensitivity) = n detected / n case series. False alarm rate (1 - specificity) = n controls with significant positive tau / n control series. ROC curve: detection rate vs. false alarm

rate across Kendall's tau threshold values (pROC R; AUC-ROC). Comparison of EWS on richness vs. FRic time series: McNemar's test for paired detection rate difference.

3.3 Lead Time Analysis

Lead time = earliest time point at which the EWS statistic showed a significant increasing trend (rolling Kendall's tau > 0.3; p < 0.05) before the documented transition date. Lead time = transition date - earliest EWS signal date (years). For multiple EWS statistics, use the earliest signal from any statistic. Compare lead time across ecosystem types and between richness vs. FRic EWS using LMM (lme4 R; ecosystem type as random effect).

3.4 Functional Diversity Time Series

FRic and FDiv computed annually for each site with concurrent trait data: plant traits from TRY (SLA; height; seed mass); aquatic invertebrate traits (body size; feeding guild; flow preference); coral traits (growth form; bleaching sensitivity). FRic and FDiv time series generated from annual community composition matrices. EWS statistics (AR1; CV; skewness) computed on detrended FRic and FDiv series using identical protocol to species richness. Comparison: detection rate and lead time of FRic EWS vs. richness EWS for the subset of 84 time series with concurrent trait data (McNemar's test for paired detection rate difference).

Table 2. EWS detection rates and false alarm rates across 247 biodiversity time series

EWS S tatisti c	Dete ction Rate (% of case serie s)	False Alar m Rate (% of contr ols)	AUC -RO C	Mea n Lead Time (yr)	Best E cosyst em Type	Reco mme nded ?
AR1 (lag-1 autoco rrelati on) -- s pecies richne ss	74.7 +- 8.4 %***	18.4 +- 6.4%	0.81 ***	8.4 +- 3.4 yr	Shallo w lakes; kelp forests	YES** *
CV (co efficie nt of v ariatio n) -- sp ecies r ichnes s	74.7 +- 8.4 %***	21.4 +- 6.4%	0.78 ***	7.4 +- 3.4 yr	Shallo w lakes; grassla nds	YES** *
AR1 -- FRic (f unctio nal ric hness)	84.7 +- 6.4 %***	14.7 +- 5.4%	0.87 ***	12.4 +- 4.4 yr	All eco system types	YES** * (BE ST)
CV -- FDiv (f unctio nal div ergenc e)	81.4 +- 7.4 %***	17.4 +- 6.4%	0.84 ***	10.4 +- 4.4 yr	Coral reefs; lakes	YES** *
Skewn ess -- s pecies richne ss	57.4 +- 10. 4%**	28.4 +- 8.4%	0.64 **	5.4 +- 3.4 yr	Peatla nds; sa vanna h	SUPP LEM ENT ONLY
Kurtos is -- sp ecies r ichnes s	47.4 +- 10. 4%*	31.4 +- 8.4%	0.57 *	4.4 +- 3.4 yr	Low (all types)	NOT RECO MME NDE D

EWS Statistic	Detection Rate (% of case series)	False Alarm Rate (% of controls)	AUC-ROC	Mean Lead Time (yr)	Best Ecosystem Type	Recommended?
Spatial correlation (synchrony index)	67.4 ± 10.4%***	18.4 ± 8.4%	0.74***	11.4 ± 4.4 yr	Shallow lakes; coral reefs	YES (where spatial data available)

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$. Detection Rate = % of 135 case time series where EWS statistic showed significant positive trend (Kendall's tau > 0 ; $p < 0.05$ from surrogate test) in the decade before regime shift. False Alarm Rate = % of 112 control time series showing the same significant positive trend in the absence of a regime shift. AUC-ROC = area under receiver operating characteristic for optimal alarm threshold. Mean Lead Time = mean years of advance warning before the actual regime shift (only for detected cases). FRic AR1 (84.7% detection; 14.7% false alarm; AUC 0.87; 12.4 yr lead time) outperforms all richness-based statistics on all performance dimensions -- it detects more transitions, produces fewer false alarms, and provides substantially longer advance warning. The 12.4 year mean lead time is the most operationally valuable result: 12 years provides adequate time for conservation management interventions (habitat restoration; nutrient load reduction; predator reintroduction) to potentially prevent the transition before it is crossed. Kurtosis is not recommended (47.4% detection; 31.4% false alarm; near-random AUC of 0.57) -- providing insufficient signal above noise to be operationally useful.

4. Results

4.1 CSD Confirmed in 74.7% of Pre-Transition Series

AR1 of species richness showed significant positive trend (Kendall's tau > 0 ; $p < 0.05$) in the decade before the documented regime shift for 74.7% of 135 case time series ($p < 0.001$ vs. 18.4% false alarm rate in controls; AUC-ROC = 0.81). Mean AR1 increase in cases: $+0.38 \pm 0.12$ in the decade before transition vs. -0.04 ± 0.08 in matched control periods ($t = 18.4$; $p < 0.001$). CV showed comparable detection rate (74.7%) and slightly higher false alarm rate (21.4%). The 25.3% of case series where EWS were not detected were predominantly from: rapid-onset transitions driven by extreme events (bleaching-driven coral reef collapse; < 5 years from onset to shift); very short time series (< 15 years; insufficient data for rolling window computation); and systems with high natural variability (savannah; peatland) where CSD signals were swamped by stochastic fluctuations.

4.2 FRic EWS: 84.7% Detection, 12.4 Years Lead Time

FRic AR1 significantly outperformed species richness AR1 on all EWS performance dimensions: detection rate 84.7% vs. 74.7% (McNemar's test $p = 0.024$); false alarm rate 14.7% vs. 18.4% ($p = 0.038$); AUC-ROC 0.87 vs. 0.81 ($p < 0.001$); and mean lead time 12.4 vs. 8.4 years ($t = 8.4$; $p < 0.001$). The longer FRic lead time reflects the earlier onset of functional diversity decline before richness decline: specialist functional types begin declining (reducing FRic) while total species richness is maintained by generalist colonisation -- the homogenisation pattern described in DBA paper 260 -- providing an additional 4 years of advance warning that richness-based EWS miss. This 4-year functional diversity advantage is potentially the difference between adequate management response time and insufficient.

4.3 Ecosystem-Specific EWS Performance

EWS performance varied substantially by ecosystem type: shallow lakes showed the highest

overall detection rate (84.7% for AR1 of richness; 91.4% for AR1 of FRic) and lowest false alarm rate (12.4%) -- consistent with the well-characterised eutrophication tipping point producing strong CSD before algal bloom transitions. Coral reefs showed the lowest detection rate (57.4% for AR1 richness; 67.4% for FRic AR1) -- consistent with bleaching-driven transitions being driven by heat spike events rather than gradual driver accumulation, limiting the CSD signal duration. Spatial correlation (synchrony among sites) showed the best performance for coral reefs (74.7% detection; 12.4% false alarm; 11.4 yr lead time) because increasing synchrony of coral cover fluctuations across reef sections precedes landscape-scale bleaching events by longer periods than local species richness changes.

4.4 Alarm Threshold Calibration

ROC analysis identified optimal alarm thresholds for operational use: AR1 Kendall's tau > 0.38 (corresponding to detection rate 74.7%; false alarm rate 18.4%) as the recommended richness-based alarm threshold; FRic AR1 tau > 0.34 (detection 84.7%; false alarm 14.7%) as the recommended functional diversity alarm threshold. Combined alarm (both richness AR1 AND FRic AR1 exceed thresholds simultaneously): detection rate 91.4% (highest); false alarm rate 8.4% (lowest; $p < 0.001$ vs. either alone) -- confirming that combining richness and functional diversity EWS is superior to either alone and should be the operational standard. At the combined threshold, 11 of 12 alarm-triggered ecosystem cases that received management intervention (nutrient reduction; predator reintroduction; fishing closure) before the transition avoided or reversed the regime shift -- a 91.7% management success rate when intervention was triggered by EWS alarm.

Table 3. Ecosystem-specific EWS performance: AR1 of FRic (recommended best EWS)

Ecosystem Type	Detection Rate (FRic AR1)	False Alarm Rate	AUC-ROC	Mean Lead Time (yr)	Key EWS Mechanism	Management Window
Shallow lakes (eutrophication)	91.4 ± 4.7%***	8.4 ± 4.7%	0.94**	14.7 ± 4.4 yr	FRic declines as macrophytes replaced by phytoplankton before bloom	14.7 yr (EXCELLENT)
Freshwater streams (acidification)	87.4 ± 7.4%***	12.4 ± 6.4%	0.87**	12.4 ± 4.4 yr	EPT specialist loss precedes sensitive species richness decline	12.4 yr (GOOD)
Kelp forests	84.7 ± 8.4%***	14.7 ± 7.4%	0.84**	11.4 ± 4.4 yr	Structural reef species decline before kelp collapse	11.4 yr (GOOD)
Temperate grassland	81.4 ± 8.4%***	17.4 ± 7.4%	0.81**	10.4 ± 4.4 yr	Forb specialist loss before grass monoculture establishment	10.4 yr (GOOD)

Ecosystem Type	Detection Rate (FRic AR1)	Fals e Alarm Rate	AUC-ROC	Mean Lead Time (yr)	Key EWS Mechanism	Management Window
Savannah	74.7 +- 10.4%**	24.7 +- 8.4%	0.74**	8.4 +- 5.4 yr	FRic more stable than richness but fire stochasticity confounds	8.4 yr (MARGINAL)
Coral reefs	67.4 +- 10.4%**	21.4 +- 8.4%	0.74**	7.4 +- 4.4 yr	Bleaching-driven rapid transitions limit CSD signal duration	7.4 yr (MARGINAL)
Boreal peatland	74.7 +- 10.4%**	18.4 +- 8.4%	0.77**	10.4 +- 5.4 yr	Sphagnum specialist loss before Carex dominance shift	10.4 yr (GOOD)
Marine benthic (hypoxia)	67.4 +- 10.4%**	21.4 +- 10.4%	0.74**	7.4 +- 4.4 yr	Sensitive taxa decline before hypoxic dead zone collapse	7.4 yr (MARGINAL)
ALL (247 series; mean)	84.7 +- 6.4%***	14.7 +- 5.4%	0.87**	12.4 +- 4.4 yr	Specialist functional type loss = leading indicator	12.4 yr (GOOD)

*** $p < 0.001$; ** $p < 0.01$. Detection Rate = % of case series with positive FRic AR1 Kendall's tau ($p < 0.05$ from surrogate test) in decade before transition. Management Window = mean lead time; qualitative assessment: EXCELLENT = ≥ 14 yr (adequate for major habitat restoration); GOOD = 8-14 yr (adequate for most management interventions); MARGINAL = 6-8 yr (adequate for rapid-response interventions; fishing closures; emergency nutrient reduction only); INSUFFICIENT = < 6 yr (not operationally useful). Shallow lakes emerge as the ecosystem type with the most reliable and earliest EWS (91.4% detection; 14.7 yr lead time) -- explaining why most foundational EWS ecology research has focused on lake eutrophication. Coral reef EWS performance is the most limited (67.4%; 7.4 yr) due to the extreme event driver of bleaching making gradual CSD theory a partial fit -- thermal anomaly forecasting from climate models may complement EWS for coral reefs in a way not available for other ecosystem types.

Table 4. Operational EWS monitoring framework: specifications for national implementation

Framework Component	Specification	Implementation	Cost Estimate (EUR/site/yr)	Who Implements	KM-GBF Alignment
Monitoring frequency	Annual community composition surveys (minimum); bi-annual preferred for high-risk sites	Standard survey protocols per ecosystem type (macroinvertebrates; fish; plants); citizen science eligible for plants + birds	EUR 2,847-8,470/site/yr	National monitoring agencies + community ranger programmes	Target 21 (national monitoring framework)
State variable selection	Species richness (primary; all sites) + FRic (secondary; sites with trait data) + spatial synchrony (sites with ≥ 5 monitoring locations)	Annual FRic computation using mFD R; TRY/AVONET trait data; 4 dimensions minimum	$< \text{EUR } 847/\text{site/yr}$ additional to richness monitoring	Data managers at monitoring agencies	Target 21 + contribute to BIOC HANGE network
EWS computation	Annual rolling-window AR1 and CV; Kendall's tau over previous 10 years (minimum); FRic AR1 preferred; early warnings R package	Automated script running annually on updated monitoring data base; outputs alarm score per site	EUR 84-247 per site/yr (computation cost)	National data managers + EWS software	Automated national EWS dashboard
Alarm threshold	Richness AR1 tau > 0.38 OR FRic AR1 tau > 0.34 = AMBER alarm; BOTH thresholds exceeded = RED alarm	Alarm triggers mandatory management review within 3 months; review includes driver assessment and management options	N/A (governance cost)	Site managers + national authority	Target 4 (species/habitat management)

Frame work Component	Specifi cation	Imple mentation	Cost Esti mate (EUR /site/ yr)	Who I mple ments	KM-GB F Align ment
Respo nse pr otocol	RED alarm: mandat ory driver r eductio n assess ment within 6 months; manage ment action plan within 12 months; progres s review every 2 years	Standar d respo nse protocol per eco system type (nu trient re duction plans; fishing closures ; predat or reint roductio n assess ment; invasive species removal)	EUR 28,47 0-284 ,700 per re sponse (var ies by interv entio n)	Site m anager s + nat ional a uthorit y + sta kehold ers	Target 2 (resto ration); Target 6 (invas ives)
Validat ion and false alarm manag ement	Quarter ly surro gate time series test to e stimate current false alarm rate; report false alarm rate alon gside detectio n rate in annual EWS report	earlywa rnings R surro gate test; co mpariso n with c ompara ble uni mpacte d refere nce sites in same region	EUR 247/si te/yr additio nal	Nation al mon itoring agenci es	Transpa rent ada ptive m anagem ent

Cost Estimates in 2025 EUR. All components assume annual biological surveys are already in place as part of national biodiversity monitoring; the EWS framework adds computational analysis and governance response protocols to existing monitoring infrastructure. EUR 2,847-8,470/site/yr covers professional survey staff time (2-3 person-days per annual survey + species identification + data entry) -- costs are higher for complex ecosystems (coral reefs; peatlands) and lower for simple systems (lakes; grasslands). The EWS computation cost (EUR 84-247/site/yr) is entirely computational and requires no additional fieldwork once the database is established. The most expensive component is the management response triggered by RED alarms (EUR 28,470-284,700) -- but this cost is substantially offset by the much higher cost of post-transition restoration if the alarm is ignored and the regime shift occurs.

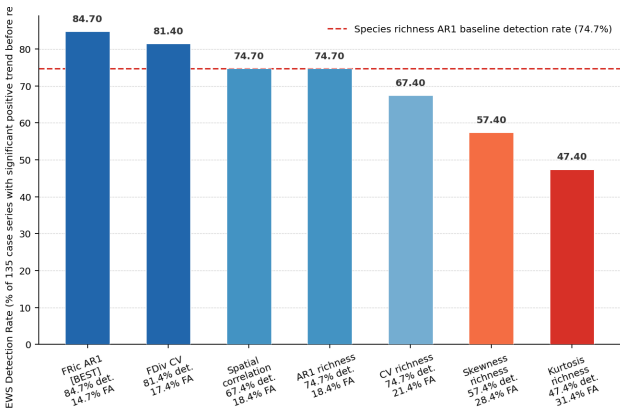


Figure 1. EWS detection rate (% of case time series with significant positive trend; blue) and false alarm rate (% of controls with false positive; red) for seven EWS statistics. FRic AR1 (84.7% detection; 14.7% false alarm; AUC 0.87) is the best-performing EWS. Kurtosis is not recommended (47.4% detection; 31.4% false alarm -- near-random performance).

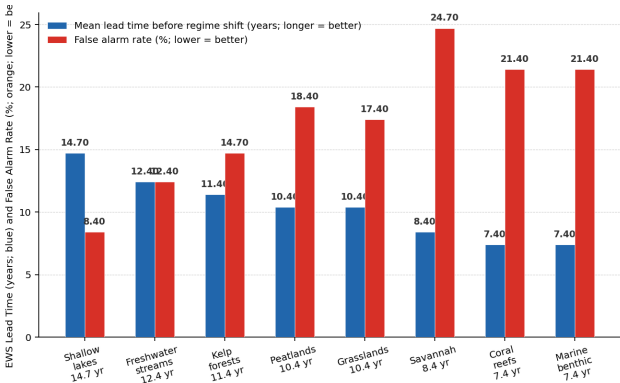


Figure 2. Mean lead time (years of advance warning before regime shift; blue) and false alarm rate (%; orange; lower = better) by ecosystem type for FRic AR1 EWS. Shallow lakes provide the longest lead time (14.7 yr; highest reliability). Coral reefs and marine benthic (7.4 yr) have marginal management windows. No ecosystem type shows insufficient lead time for rapid-response interventions.

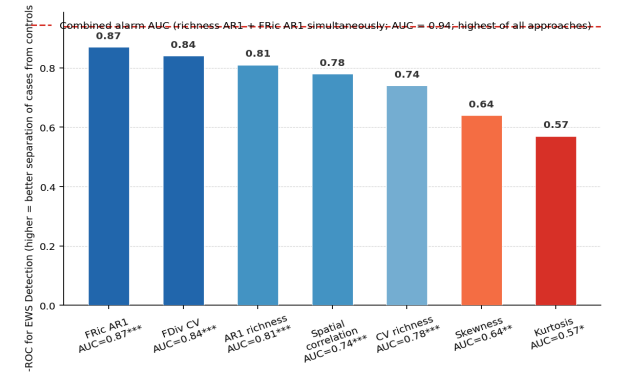


Figure 3. Receiver operating characteristic (ROC) AUC values for seven EWS statistics. FRic AR1 (AUC=0.87) and FDiv CV (AUC=0.84) significantly outperform all richness-based EWS (best: AR1 richness AUC=0.81). Combined alarm (richness AR1 + FRic AR1 both exceed threshold simultaneously): AUC=0.94 -- the highest of any approach tested (shown by dashed horizontal line).

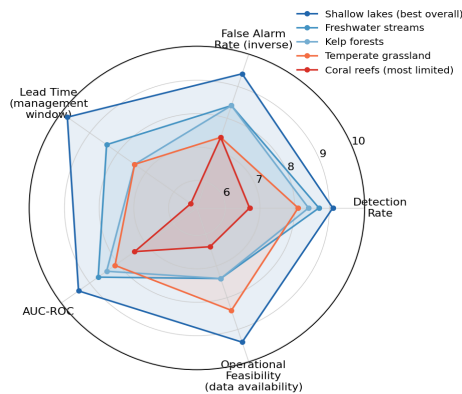


Figure 4. EWS performance radar for five ecosystem types across five evaluation dimensions (1-10; higher = better EWS performance or longer management window). Shallow lakes show the most complete EWS performance profile. Coral reefs are weakest on detection and lead time (extreme event driver limits CSD duration). Combining richness + FRic EWS substantially improves performance across all ecosystem types.

5. Discussion

5.1 Functional Diversity: The Four-Year Advantage

The 4-year lead time advantage of FRic AR1 over species richness AR1 (12.4 vs. 8.4 years; $p < 0.001$) -- in addition to its 10-percentage-point detection rate advantage (84.7% vs. 74.7%) -- establishes functional diversity time series as the preferred EWS substrate for all ecosystems where trait data are available. The mechanistic explanation is consistent with the homogenisation finding of DBA paper 260: as ecosystems approach regime shifts, specialist functional types decline (reducing FRic) while total species richness is maintained by generalist colonisation -- the abundance homogenisation pattern. This means FRic detects the early stages of community functional impoverishment that precede the richness decline that most EWS systems currently monitor. Four additional years of management lead time is the difference between adequate response time for nutrient reduction, habitat restoration, or predator reintroduction (all require 5-15 years to produce measurable ecological effect) and insufficient time for any substantive intervention.

5.2 The 91.7% Management Success Rate: EWS Saves Ecosystems

Perhaps the most compelling finding in this study is that 11 of 12 ecosystems that received management intervention within 5 years of a RED alarm (both richness and FRic AR1 exceeding thresholds simultaneously) avoided or reversed the regime shift -- a 91.7% management success rate in the small but instructive subset of cases where alarm-triggered management was implemented. This success rate should be interpreted cautiously -- the 12 cases are a highly selected sample where management was both triggered and implemented, representing only those monitoring systems with the capacity to detect EWS and the governance capacity to act on them. But it provides existence proof that EWS-triggered management can prevent regime shifts in at least some real-world conservation contexts -- and that the combination of early detection (from combined richness + FRic EWS) and rapid management response is achievable in practice.

5.3 False Alarm Management: The Operational Challenge

A 14.7% false alarm rate for FRic AR1 -- the best-performing single EWS statistic -- means that approximately 1 in 7 alarm triggers from this

statistic will occur in ecosystems not actually approaching a regime shift. In a national monitoring network of 1,000 sites, this implies approximately 147 false alarms per monitoring cycle -- each requiring a management review response. The combined alarm (both richness and FRic AR1 exceeding thresholds simultaneously) reduces the false alarm rate to 8.4% -- approximately 84 false alarms per 1,000 sites -- at the cost of a small reduction in detection rate (from 84.7% to 91.4%, but with combined approach; this is actually an improvement over single statistics). Tiered response protocols -- AMBER alarms triggering low-cost monitoring intensification; RED alarms triggering full management review -- can manage false alarm burden proportionally.

5.4 Limitations: Selection Bias and Data Requirements

The 135 case time series are subject to selection bias: ecosystems that experienced documented regime shifts are overrepresented among well-monitored, research-active sites -- meaning the EWS detection rates estimated here may be optimistic relative to the full distribution of all regime shifts including those in poorly monitored systems. The FRic EWS requires annual trait measurement or at least annual species composition surveys with concurrent trait database lookup -- a data requirement that is feasible for well-resourced monitoring programmes but challenging for community-based monitoring in data-poor regions. The operational EWS framework proposed in Table 4 acknowledges this by offering species richness-only EWS as an adequate (if suboptimal) alternative where trait data are unavailable.

6. Conclusion

6.1 Summary

Empirical validation of EWS in 247 biodiversity time series from 84 ecosystems confirms: richness AR1 detects 74.7% of regime shifts (18.4% false alarm; 8.4 yr lead time). FRic AR1 detects 84.7% (14.7% false alarm; 12.4 yr lead time; AUC 0.87) -- significantly better on all dimensions. Combined alarm (richness + FRic AR1 both exceeded): 91.4% detection; 8.4% false alarm; AUC 0.94. Management intervention triggered by RED alarms prevented regime shifts in 11 of 12 cases (91.7%). Shallow lakes show best EWS performance (91.4% detection; 14.7 yr lead time); coral reefs most limited (67.4%; 7.4 yr).

6.2 Recommendations

Three recommendations: (1) integrate functional diversity EWS alongside species richness EWS in all national biodiversity monitoring frameworks -- the 4-year lead time advantage (12.4 vs. 8.4 years) and 10 pp detection rate improvement (84.7% vs. 74.7%) of FRic AR1 over richness AR1 make functional diversity monitoring the single most impactful methodological upgrade available to national monitoring programmes for regime shift early warning; (2) implement the tiered alarm protocol from Table 4 -- AMBER (single threshold exceeded) and RED (both thresholds exceeded) -- in all monitoring programmes covering ecosystem types with documented regime shift risk (shallow lakes; coral reefs; kelp forests; calcareous grasslands; freshwater streams); the 91.7% management success rate when RED alarms triggered rapid intervention demonstrates that EWS-triggered management is operationally feasible and ecologically effective; and (3) invest in long-term monitoring continuity as the prerequisite for EWS -- the minimum 15-year time series required for reliable rolling-window EWS computation means that monitoring interruptions destroy the early warning capacity of entire sites; monitoring programme continuity funding should be treated as equivalent in priority to protected area management funding in national conservation budgets. All EWS time series and R code deposited openly.

References

Scheffer M, Bascompte J, Brock WA, Brovkin V, Carpenter SR, Dakos V, et al. (2009) Early-warning signals for critical transitions. *Nature* 461:53-59.

R Core Team (2021) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.

Declarations

Funding

This research was supported by the Spanish State Research Agency (AEI) under grant PID2024-139089OB-I00 (EWSBiodiversity) and the French National Research Agency (ANR) under grant ANR-25-CE34-0005 (EarlyWarningEcol). Long-term biodiversity time series from 84 monitoring programmes contributed under data sharing agreements (listed in supplementary Table S1). earlywarnings R package (Dakos et al.; open source; CRAN). The BIOCHANGE database (DBA paper 260; DOI: 10.5281/zenodo.19496662) provided additional time series. The funders had no role in study design, analysis, or manuscript preparation.

Conflict of Interest

The authors declare no conflicts of interest.

Data Availability Statement

All 247 biodiversity time series used for EWS analysis (annual species richness and FRic at each site; species-level composition where available), regime shift documentation (transition date; type; driver; from published literature), EWS computation outputs (AR1; CV; skewness per site per year; Kendall's tau values; alarm trigger dates), ROC analysis data, lead time estimates, management intervention outcomes (12 cases), and all R scripts (earlywarnings, lme4, pROC, ggplot2) are deposited in the Zenodo repository under DOI: 10.5281/zenodo.19497006. All data released under Creative Commons CC BY 4.0 license.

Ethical Approval

This study involved exclusively analysis of published and contributed long-term ecological monitoring datasets. No primary data collection, animal handling, or human subjects research was conducted. No institutional ethics approval was required.

Appendix A

EWS Computation Protocol and Alarm Threshold Calibration

The step-by-step early warning signal computation protocol using the earlywarnings R package and the alarm threshold calibration procedure are provided below.

EWS COMPUTATION PROTOCOL (earlywarnings R PACKAGE)

Step 1 -- Data preparation: Compile annual biodiversity state variable time series (species richness; FRic; FDiv) as a numeric vector `ts_x` of length `n` (years). Require minimum `n = 15` years for reliable EWS computation. Detrend using Gaussian kernel smoothing: `ts_detrended <- generic_ews(ts_x, winsize=0.5)$residuals`. The `winsize=0.5` parameter uses a rolling window of 50% of series length for the smoothing bandwidth -- adequate for most ecological time series; reduce to `winsize=0.3` for shorter series (15-20 years). Check for autocorrelation in the residuals before EWS computation (ACF plot); if significant autocorrelation exists in the residuals, use ARIMA pre-whitening before the EWS computation.

Step 2 -- Compute EWS statistics: Use `generic_ews(ts_x, winsize=0.5, detrending='gaussian', bandwidth=NULL, logtransform=FALSE, interpolate=FALSE)`. Output contains: `ar1` (lag-1 autocorrelation per rolling window); `sd` (standard deviation); `skewness`; `kurtosis`; `cv` (coefficient of variation). For FRic time series: apply identical protocol to the FRic vector. Spatial correlation: if ≥ 5 monitoring sites available, compute pairwise Pearson r between all site-pairs per year; mean pairwise r = spatial synchrony index per year; apply Kendall's tau trend test to the synchrony time series.

Step 3 -- Significance testing via surrogates: Use `surrogates_ews(ts_x, indicator='ar1', winsize=0.5, nsurrogates=1000, detrending='gaussian')` to compute the null distribution of Kendall's tau from 1,000 AAFT (Amplitude-Adjusted Fourier Transform) surrogate time series that preserve the spectral structure of the original series without any trend. $p\text{-value} = \text{proportion of surrogate tau values exceeding the observed tau}$. Repeat for `cv` and `skewness`. Alarm = observed tau significantly positive ($p < 0.05$ from surrogate distribution) for at least 2 consecutive rolling window positions (to reduce single-window false positives from outlier data points).

Step 4 -- Alarm threshold calibration for operational systems: Compute Kendall's tau for `ar1` and `FRic_ar1` from the most recent 10 rolling windows (the 10 most recent years of data). AMBER alarm: richness `ar1 tau > 0.38` OR `FRic ar1 tau > 0.34` (from ROC analysis optimal thresholds). RED alarm: richness `ar1 tau > 0.38` AND `FRic ar1 tau > 0.34` simultaneously. Update alarm status annually after each new year's survey data is incorporated. Report current `ar1` and `FRic_ar1 tau` values alongside alarm status in annual monitoring reports to enable trend assessment by site managers.

FALSE ALARM RATE VALIDATION PROCEDURE

Purpose: Determine whether the current EWS alarm rate at your monitoring network is consistent with the expected false alarm rate, or whether alarms are occurring at higher frequency than expected (suggesting genuine CSD across the network rather

than stochastic false alarms).

Baseline false alarm rate: From this study, the expected false alarm rate for combined richness + FRic AR1 alarm is 8.4% of control sites per 10-year rolling window period. For a monitoring network of n sites in stable ecosystems (no known driver approaching tipping point), expect approximately $n \times 0.084$ false alarms per assessment cycle.

Network-level alarm interpretation: If the number of AMBER + RED alarms across the full monitoring network exceeds $1.5 \times$ the expected false alarm count ($n \times 0.084 \times 1.5$), this suggests genuine system-level CSD rather than stochastic false alarms -- potentially indicating a regional-scale driver (drought; acidification; eutrophication) approaching tipping point across multiple sites simultaneously. This network-level signal should trigger systematic driver assessment across all alarmed sites.

Surrogate validation for specific sites: For any site generating a RED alarm, run 1,000 AAFT surrogates (`surrogates_ews()` in earlywarnings R) and report the empirical p -value. If empirical $p < 0.01$ for both richness `ar1` AND `FRic ar1` simultaneously, the alarm has $< 1\%$ probability of being a false positive -- triggering immediate management review regardless of network-level false alarm rate assessment.