Stability metrics are predictable across data qualities but sensitive to community size.

Abstract

Modern biodiversity monitoring is generating increasingly multidimensional representations of wildlife populations and ecosystems. It is therefore appealing for conservation and environmental governance to combine that information into single measure of ecosystem or population health. Stability is a desirable feature of ecosystems that supports this aim, measured through resistance, recovery, and variability. The Jacobian matrix is a common characteristic used to identify the resistance and resilience of deterministic mathematical systems and whilst historically it has been challenging to estimate from empirical data, recent work has proposed a suite of metrics capable of reconstructing it for a real-world community using time series data. Here we assess the robustness of three Jacobian metrics and two variability estimating stability metrics to varying time series lengths and data qualities based on that seen in real-world wildlife time series. Using Lotka–Volterra equations, we generate short time series (to match global biodiversity datasets such as the Living Planet Index and BIOTIME) and sampling error corruptions (to mimic varying search efforts) to validate metric performance in empirical data. The robustness of all stability metrics improved with time series length and sampling error in the anticipated manner. However, number of species dramatically altered metric capability, with larger communities decreasing the reliability of stability metric trends. Overall, stability metrics behave predictably across realistic data corruptions. Generic stability estimation is therefore possible from abundance time series alone, and we suggest that, given the increasing availability of multivariate community data, focussing on Jacobian estimates is a plausible ecosystem condition indicator.

Introduction

Biodiversity loss is a major concern for the maintenance of ecosystem stability and functioning (Cardinale et al., 2012; Oliver et al., 2015). Despite this, species extinctions are accelerating (Ceballos et al., 2015) with resulting change in ecosystem services, exploitable resources, and habitat cover being sufficiently impactful (Traill et al., 2011; Tylianakis et al., 2008) that biodiversity policy targets have become commonplace (United Nations, 2020). The conservation and political decisions behind the selection of these targets is dependent on sufficiently high quality data and reliable indicators of biodiversity change (Dale & Beyeler, 2001). Such biodiversity indicators are ultimately used to quantify both the scale of impacts and monitor mitigation actions – i.e. have the desired outcomes been achieved (Donatti et al., 2020) – in an "indicator-policy cycle" (Nicholson et al., 2012) to ensure appropriate policy decisions during a period of competing priorities for decision makers.

The Living Planet Index (LPI - Loh et al., 2005) has emerged alongside BIOTIME (Dornelas et al., 2018), the Global Population Dynamics Database (Inchausti & Halley, 2001) and the IUCN Red list (IUCN, 2022) as the focal datasets and descriptive instruments for global biodiversity trends. They have influenced international (United Nations, 1992, 2020), national (Environment and Climate Change Canada, 2019), and regional (Brotons et al., 2020) environment policy, as well as scientific discourse (Dornelas et al., 2019; Gonzalez et al., 2016; Johnson et al., 2024; Leung et al., 2020; Loreau et al., 2022; McRae et al., 2017). It is therefore undeniable the importance of these datasets for current and future biodiversity monitoring especially given the relative paucity of long-term and reliable counts of species globally (Hughes et al., 2017; McRae et al., 2017). Consequently, the available time series have driven a primary focus on abundance changes by biodiversity indicators in absolute or relative terms (i.e. species richness or diversity indices), despite the acknowledged sampling

errors, geographical biases, and many short time series present (Gonzalez et al., 2016; Leung et al., 2020). These concerns generate uncertain estimates of global biodiversity change (Johnson et al., 2024; Wauchope et al., 2019) which hamper the indicator-policy cycle.

Such a focus on abundance also potentially overlooks other signals expressed by stressed populations (Cerini et al., 2023) which are likely mechanistically important for abundance decline and can provide earlier and more robust indications of biodiversity loss. Unfortunately, for many practitioners, abundance data is all that is available, leading to alternative suggestions to consider the stability of the system (Capdevila et al., 2022; Holling, 1973) rather than focussing on raw abundance trends. This is achieved by reconciling the interacting nodes/species into a singular measure of population/ecosystem health (Dakos et al., 2015; Ushio et al., 2018) – an indicator of system stability dependent on the response of the system rather than solely its species diversity. Here we define stability following Pimm (1984) as the capacity to 'return to the [system's] initial equilibrium [or state] following [it] being perturbed from it'. However, measuring ecosystem stability from empirical data is challenging (Capdevila et al., 2022; Donohue et al., 2013, 2016) with an ever expanding diversity of possible metrics and indices proposed (Kéfi et al., 2019). Kéfi et al. (2019) and Donohue et al. (2013) specifically highlight that stability indicators are rarely considered in parallel, across neither unidimensional (i.e. populations) nor multidimensional (i.e. communities) scales. Empirical stability research should therefore attempt to reconcile such differences in scale and dimensionality. In addition, Kéfi et al. (2019) also suggest that any metric must be rooted in theory and be consistent across empirical and theoretical studies. For example, the dominant eigenvalue of the Jacobian is common in theoretical studies (Baruah et al., 2022; Dakos, 2018; Strogatz, 2015), but has rarely been reported in real-world studies (Kéfi et al., 2019) due to the challenge of defining empirical Jacobians.

Computational advancements in the last decade have provided a novel toolbox for estimating Jacobians and species interaction matrices from observational data (Ahmad et al., 2016; Sugihara et al., 2012; Williamson & Lenton, 2015). These indicators are grounded in dynamical system theory (Strogatz, 2015) and have the potential to quantify the resilience dimension of stability under strong equilibrium but also under more complex and chaotic dynamics (Lyapunov, 1992). This is obviously attractive for biodiversity monitoring as loss of resilience/stability is indicative of oncoming decline (Dakos et al., 2015) even if the system is away from equilibrium (which is likely the case for ecosystems experiencing multiple stressors (Hastings et al., 2018)). However, it is unclear how stability metrics behave across data qualities and ecosystem community structure to warrant their immediate usage in real-world data. Similar ambiguities have influenced the practicality of related population collapse indicators (a.k.a. Early Warning Signals – EWSs) with good performance reported in idealised simulated studies (Clements et al., 2015; Dakos, 2018) but failures in observational time series (Burthe et al., 2016; Zhang, 2020). There is consequently a need to test the behaviour of the stability metric toolbox in controlled scenarios matching the data qualities present in biodiversity datasets such as the LPI, and ensure their behaviour is sufficiently reliable prior to their potential usage in the indicator-policy cycle.

In this work, we use simulated multispecies ecological communities to test the influence of time series length and data corruption upon five stability metrics in a system where the true dynamics are known. These stability metrics included a multivariate index of variability (a measure of variability - (Brock & Carpenter, 2006)), Fisher information (a measure of variability and resistance -(Ahmad et al., 2016; Karunanithi et al., 2008)), and three Jacobian matrix estimators (measures of resistance and resilience - (Grziwotz et al., 2023; Ushio et al.,

2018; Williamson & Lenton, 2015)). We specifically estimated the trend of each metric in 5, 15 and 25 species communities across different scales (populations vs communities) and show that while metric reliability generally increases with data quality, metrics displayed weak but correct trends even in low data qualities. However, the choice of species contributing to the metric dramatically influenced reliability leading to recommendations that users must guide their species selection with ecological knowledge.

Methods

Simulation framework

Community dynamics were simulated using an established multi-species generalised Lotka-Volterra competitive model (Hofbauer & Sigmund, 1998), containing a control parameter that can drive one or more species to extinction. Specifically, this model simulates the appearance of an invasive species (Equation 1 - (Dakos, 2018)) in an established competitive community and is written as:

$$\frac{\mathrm{d}N_i}{\mathrm{d}t} = r_i N_i(t)\left[1 - \frac{\sum_{j=1}^{n} \alpha_{ij} N_j(t)}{f_i(t)} + \mu_i\right] + \varepsilon_i(t)$$

$$f_i(t) = K_i(1 + \eta_i C)$$

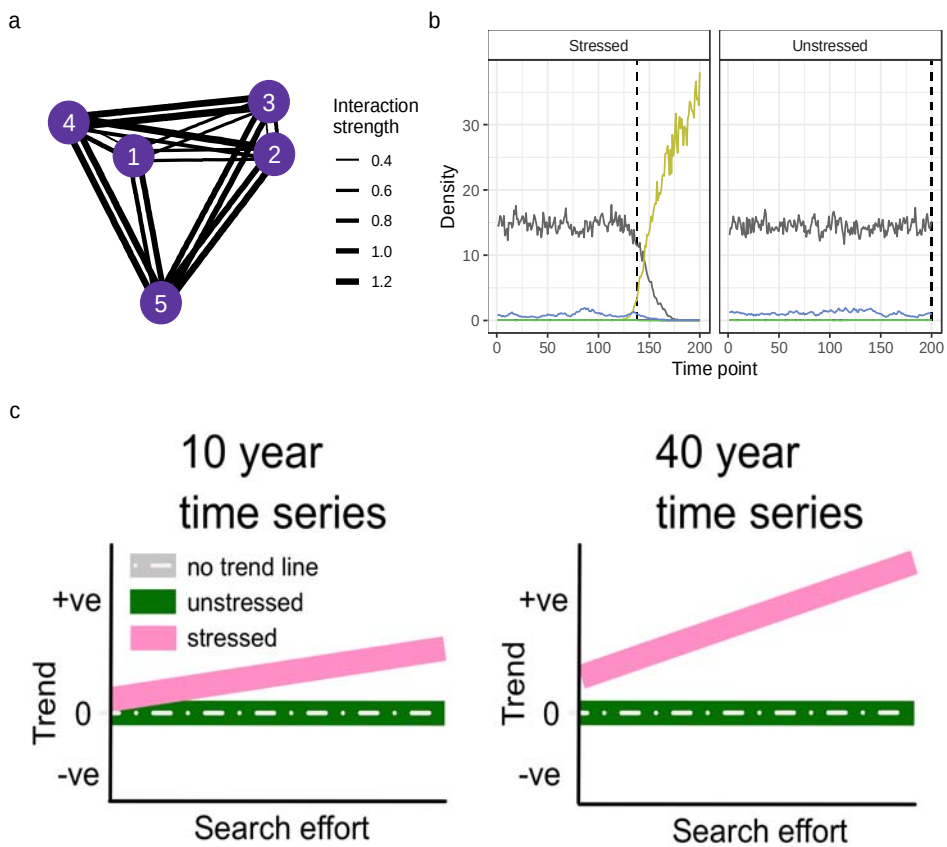
Equation 1

where i and j are the focal and interacting species identities respectively. N_i is the resulting population density of species i, r_i is i's reproduction rate, α_{ij} is the interaction coefficient between species i and j (or the per capita effect of species j on i), f_i is the species-specific invasion function with C the control parameter, and ε is a white noise process with mean 0

and variance 0.1. Symbols K_i, μ_i, and η_i represent the initial carrying capacity, immigration term, and species-specific response to the control parameter C for species i respectively.

The community parameters were parameterised followed Dakos (2018). Specifically, in this competitive community interaction matrix α_{ij} elements were drawn from a uniform distribution 0 – 1.5 with the diagonal set to 1 to ensure intraspecific competition. Reproduction rate r_i, carrying capacity K_i, and sensitivity to the control parameter η_i were sampled from the uniform distributions 0.9 – 1.1, 5 – 15, and 0 – 1 respectively. The immigration rate μ_i was set as 0.01 for all species.

Figure 1. a) Exemplary trophic structure and intraspecific interactions for a five species community. Numbers represent species, solid paths are species interactions, and the width of the path indicates the strength of interaction. b) A simulated community across stressed and unstressed conditions. The estimated inflection point is indicated by a vertical dashed line. c) Hypothesised behaviour of stability metrics through time (trend) across time series length and search effort. All metrics (excluding Fisher information) are expected to increase through time in increasingly stressed communities and remain constant in unstressed. Increasing time series length and search effort is anticipated to improve the reliability and strength of relationship. Fisher information is expected to display negative trends. A dashed zero trend line is provided for reference.



Stable communities were then identified by estimating the community densities at equilibrium (N^*) before calculating the Jacobian matrix J using the equation $J_{ij} = \alpha_{ij}N_i^*$ (May 1974) and its eigenvalues. When the dominant real eigenvalue of J is positive, the system is locally unstable, while when negative and close to zero, its return rate to equilibrium was

sufficiently slow that we rejected any community with a dominant eigenvalue greater than -0.05. We then randomly selected 25 stable communities with unique interaction matrices.

Each community was stochastically simulated using the Euler-Maruyama method at dt = 1, for 100 timesteps with a control parameter value of 0. In the following period between t100 and t200, the control parameter value was incrementally increased and then held at the maximum value of 5.0 (Dakos, 2018) from t200 to t300. For simplicity, only one species was targeted as an invasive species as such species typically appear individually (Doherty et al., 2016), though multiple can successively invade (O'Loughlin & Green, 2017). The invading species within each community was chosen at random to generate a diversity of community dynamics. Each community was simulated 25 times repeated across stressed and unstressed regimes resulting in 1250 timeseries per community size (2 stress regimes, 25 communities, 25 stochastic runs). Simulations were run using the '*DifferentialEquations*' package (Rackauckas & Nie 2017) in the Julia language (Bezanson *et al.* 2017).

To clarify, the purpose of these simulations is to create a library of stressed community dynamics where the true fate is known for indicator testing. The choice of this simulation framework ensures that the shape and trends of the community dynamics can be both linear and/or non-linear (Dakos, 2018) akin to observed wildlife time series (Johnson et al., 2024; Ledger et al., 2023), and so are appropriate to challenge these proposed stability indicators.

Data post-processing

Each simulated time series was then altered to mimic the imperfect sampling and short time series found in the LPI and BIOTIME. First, we identified the time point of fastest change in the community's first principal component using the bisection method (Christopoulos, 2016)

provided by the *inflection* R package (v.1.3.6). This method simply estimates the first derivative between successive points and searches for the largest derivative. Identifying these changes in ensures that the community is stressed but no species has been yet driven to collapse (Figure 1b), and allows us to identify practical stability metric behaviour over timescales appropriate for intervention (Davidson et al., 2018; Hastings, 2016). Consequently, each time series was truncated to 10, 25, 40, 55 and 70 timepoints prior to the estimated inflection point, matching the range of time series lengths contributing to the LPI/BIOTIME. Each truncated time series was then subjected to simulated searches ranging from 10 to 100% effort in 10% increments. To achieve this, we converted the simulated continuous species biomass data to counts, to represent natural animal populations, by multiplying the simulated data by 100 and rounding to the nearest integer. We then drew 'observed' species counts from a binomial distribution with probability equal to the search effort (Clements et al., 2015), which mimics both imperfect spatial sampling and detection. Resultantly, each initial simulation was reanalysed for each time series length and search effort combination. All post processing, metric calculation and analysis was performed in the R language v.4.3.0 (R Core Team, 2022).

Stability metrics

Five stability metrics were calculated for each post-processed simulation using the R package EWSmethods (v.1.2.0) (O'Brien et al., 2023). These metrics included Fisher information (Karunanithi et al., 2008), a multivariate index of variability (Brock & Carpenter, 2006), two S-map estimated Jacobian indices (Grziwotz et al., 2023; Ushio et al., 2018) and a Jacobian index estimated from multivariate autoregressive models (Williamson & Lenton, 2015). All metrics were calculated in a rolling window 50% the total time series length following the typical practice for generic indicators (Dakos et al., 2012; Lenton et al., 2012).

Fisher information (FI) quantifies indeterminacy or the quantity of information observed data contain on an underlying but unmeasured parameter (Ahmad et al., 2016). In practice FI is calculated by comparing changes in the probability of detecting different states relative to small inherent fluctuations in the time series. Typically, these fluctuations are defined by the variance of each time series in a reference period (Karunanithi et al., 2008). FI is therefore calculated in a rolling window along the time series, with the variability of each window compared to that of the previous. If the variability is greater than that of the reference variability, then the windows are binned into separate states. Here we defined this reference variability as the standard deviation of each species across the entire time series (Karunanithi et al., 2008). FI is consequently expected to decrease along with stability.

The multivariate index of variability (mvi) is the square root of the species covariance matrix's dominant eigenvalue and represents a multivariate equivalent to the univariate "variance" early warning signal (Brock & Carpenter, 2006). It is therefore expected to increase as stability decreases.

The two S-map Jacobian indices are derived from a state space reconstructed via empirical dynamic modelling (Park et al., 2021; Sugihara et al., 2012). In brief, both indices estimate the Jacobian of the community by comparing the time delayed relationships between each representative time series and predicting out from the reconstructed state-space (Ushio et al., 2018). The coefficients of the local linear model performing this out-prediction therefore represent the estimated α_{ij} of the community. All species data were scaled to mean zero and unit variance prior to metric estimation, following Chang et al (2017), to ensure all species

are of equivalent magnitude to maximise the accuracy of the S-map estimated interaction strengths.

The multivariate form of this Jacobian index (multiJI) compares all pair wise interactions between species in the community and constructs the state space by delay embedding up to *E*, where *E* is the number of species. The metric is therefore sensitive to the choice and number of variables (i.e species here) contributing to the S-map calculation (Medeiros & Saavedra, 2023; Ushio et al., 2018), and limits the minimum number of time points to *E*. In this study, as all species are contributing to system stability, all species are appropriate for inclusion. However, if a species count is constant for the duration of a window, then S-map reconstruction is impossible and therefore we excluded any species which were unchanging for the duration of the time series. This is possible in our simulations when a species either goes extinct prior to the shortened time series assessed, or low search efforts did not encounter species with small populations. Similarly, as time series length was often less than *E* in the specious communities (e.g. 10 timepoints but $E = 15$), for these communities we randomly selected species to contribute to multiJI to test this additional processing error on the capability of multiJI to accurately characterise stability change. multiJI is expected to increase as stability decreases with a value greater than 1 indicating instability, whereas a value below 1 indicates stability.

The univariate form of Jacobian index is calculated for each species independently by delay embedding the time series against lagged versions of itself (Grziwotz et al., 2023). The resulting α_{ij}/local linear model coefficients are therefore all 0, excluding the lower-off diagonal (all -1), whereas the first row elements become the coefficients for the focal species. Consequently, the system's behaviour can be approximated in the absence of complete

information and escapes concerns regarding the inclusion of variables/species (Grziwotz et al., 2023; Medeiros et al., 2022; Ushio et al., 2018). Here we set the embedding dimension *E* to 1 (1 "year") and the time lag τ to 10% the length of the time series following the coded examples of Grziwotz *et al*. (2023). To allow comparability with the multivariate stability metrics, we then averaged uniJI estimates across all species (mean_uniJI) and extracted the maximum index value observed (max_uniJI), with both metrics also expected to increase in parallel with stressor value similar to mvi and multiJI.

Finally, the multivariate autoregressive Jacobian index (multiAR) fits autoregressive models between multiple time series and uses the generated autocorrelation matrix's eigenvalues to estimate the system's Jacobian dominant eigenvalue (Williamson & Lenton, 2015). The eigenvalues of the two matrices' real components are related by the equation:

$$\lambda_j = \frac{1}{\Delta t} \log \ |a_j|$$

where λ_j is the Jacobian eigenvalues, a_j is the eigenvalues of the autocorrelation matrix and Δt is the time step of the data generating process. As Δt is typically unknown for empirical data, we have assumed it is equal to 1. multiAR is therefore conceptually similar to multiJI but exploits lag-1 autoregressive models (Hasselmann, 1988) rather than empirical dynamical modelling (Sugihara et al., 2012). Due to this similarity, species selection and data processing was identical to multiJI, with the same expected behaviour; an increase in value with stress.

Stability metric trends

All five metrics are expected to express directional change when the system is linearly forced/stressed but remain approximately constant when unstressed (Figure 1c). We therefore assessed the linear trend of each metric in a Bayesian mixed effects model as an alternative to the Kendall tau correlation comparisons typically applied to such indicator comparisons

(Baruah et al., 2022; Clements et al., 2015; Dutta et al., 2018). Comparing stability metric ability in this way allows us to extract representative global trends of the stochastically simulated communities while accounting for autocorrelation and shared features between simulations (Johnson et al., 2024). Each metric's trend across time series lengths and search efforts was consequently analysed independently using Laplace approximation in INLA (v.23.04.24) (Rue et al., 2009) due to the quantity of data. Interactions were modelled between time (numeric, scaled between 0 and 1 to allow comparability across truncated time series lengths), whether the simulation represents a stressed or unstressed model (factor, two levels), time series length (numeric, divided by 10 to be of equal magnitude as search effort) and search effort (numeric) as fixed effects. Random intercepts were also included for community identifier (factor, 30 levels) along with an autocorrelation function to account for serial dependence within a simulation, and random slopes for each community as each is expected to display the same trend. Separate models were fit for each community size (5, 15, or 25 species) due to computational constraints. Please note that mvi is modelled on the log scale due to sudden non-linearities known to occur in this metric's trend (Brock & Carpenter, 2006). The resulting model structure was:

$$y_{ij} \sim Normal(\mu_{ij})$$

$$\mu_{ij} = \mathrm{X}\beta + \mathrm{Z}\gamma_{ijk}$$

$$\gamma_{ijk} = \rho\epsilon_{ijk,t-1} + b_{0j} + b_{1j}time_{ijk} + \epsilon_{ikt}$$

where y_{ij} is the predicted metric value for the *i*th data point in the *j*th community and *k*th simulation with mean μ_{ij}. β represents the coefficients for time, stress, time series length, search effort and their interactions, b_{0j} the random intercepts, b_{1j} the random slope

coefficients, ρ the correlation between temporally related residuals (ε), and t the time point of i. X and Z are design matrices for fixed and random effect variables respectively.

We set the weakly informative priors following Rue et al. (2009)'s suggestions:

$$\beta \sim Normal(0,1)$$

$$b_{0j} \sim Normal(0, \sigma_j)$$

$$b_{1j} \sim Normal(0, \sigma_j)$$

$$\sigma_j \sim lognormal(0,1)$$

Jacobian threshold transgression

As the Jacobian indices can also signal unstable periods when greater then 1 (for the S-map derived methods) or 0 (for the autoregressive method), we also modelled the capability of these metrics to correctly classify stressed simulations using a binomial mixed effects model. As above, we included time series length, search effort and model stress as fixed effects, with random intercepts for community identifier, but also included community size (numeric) as a fixed effect with the response variable whether the threshold of 1/0 was crossed at any point in the time series, rather than the raw metric value. The model structure was therefore:

$$y_{ij} \sim Binomial(\mu_{ij})$$

$$\mu_{ij} = \mathrm{X}\beta + \mathrm{Z}\gamma_{ij}$$

$$\gamma_{ij} = b_{0j} + \epsilon_{ij}$$

with priors maintained as above.

<u>Results</u>

Metric trends

All metrics display their anticipated behaviour through time and are influenced by both the length of time series and quality of search effort (Figures 2-4, Figures S1-S3). In stressed 5 species communities (Figure 2, Tables S1-S5), increasing time series length improves the strength of trend estimate in all metrics for a given search effort level. However, increasing search effort alters trend differentially between metrics. For example, multiJI's and multiAR's trend is weakened with improved search efforts while the remainder of metrics all strengthen. Absolute strength of trend was generally weak in each metric, other than FI, owing to the influence of stochasticity in each simulation altering the timing of metric change between simulations and non-linearities in metric trend (Figures S1-S3).

Similarly, the conservative hierarchical and autoregressive modelling approach balloons uncertainty as identified in other work modelling global trends from multiple local trends (Johnson et al., 2024). In the unstressed communities, the S-map Jacobian based metrics behave as expected across time series and search efforts (approximately 0 trend).

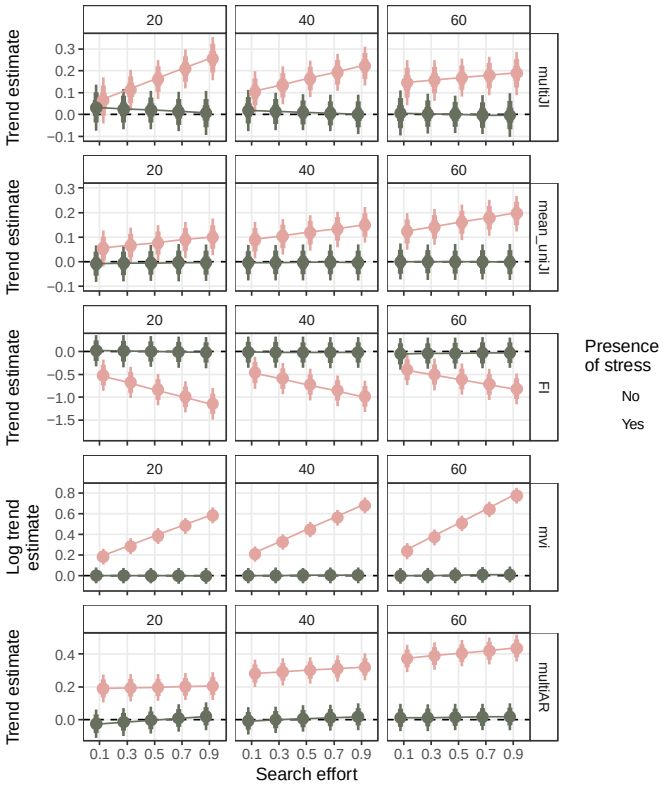


Figure 2. Posterior estimated slopes/trends of each stability metric (rows) through time, across truncated time series lengths (columns) and search efforts in 5 species communities. Dark coloured estimates represent trends in unstressed communities, while light coloured estimates are those in stressed communities. The reported values are the posterior density median values (circles), with 80% (thickest bars) and 95% (thinnest bars) credible intervals.

In the 15 species communities (Figure 3, Tables S6-S10), all metrics other than multiJI and multiAR behaved as anticipated across stressed vs unstressed communities, time series lengths and search efforts: each metric's trend was approximately zero in the unstressed communities but strengthened with time series length and search effort in the stressed. Both multiJI and multiAR were unreliable in shorter, unstressed communities though became more reliable as time series length increased. multiJI behaviour in stressed communities was stable

across time series length and search efforts and was primarily compromised by unexpectedly positive trends in unstressed communities with shorter time series.

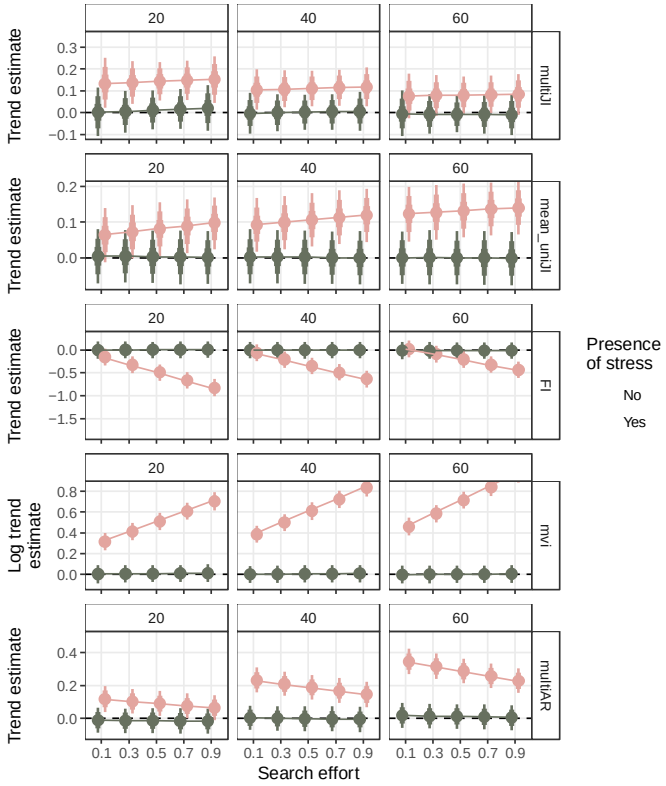


Figure 3. Posterior estimated slopes/trends of each stability metric (rows) through time, across truncated time series lengths (columns) and search efforts in 15 species communities. Dark coloured estimates represent trends in unstressed communities, while light coloured estimates are those in stressed communities. The reported values are the posterior density median values (circles), with 80% (thickest bars) and 95% (thinnest bars) credible intervals.

Finally, all metrics perform less reliably in 25 species communities (Figure 4, Tables S11-S15) relative to the smaller communities. multiJI and multiAR are unable to distinguish stressed from unstressed communities regardless of time series length or search effort while mean_uniJI only distinguishes stressed from unstressed in the longest and best searched

communities. FI and mvi both perform as anticipated across increasing time series lengths and search efforts but with weaker trends than those observed in the smaller communities.

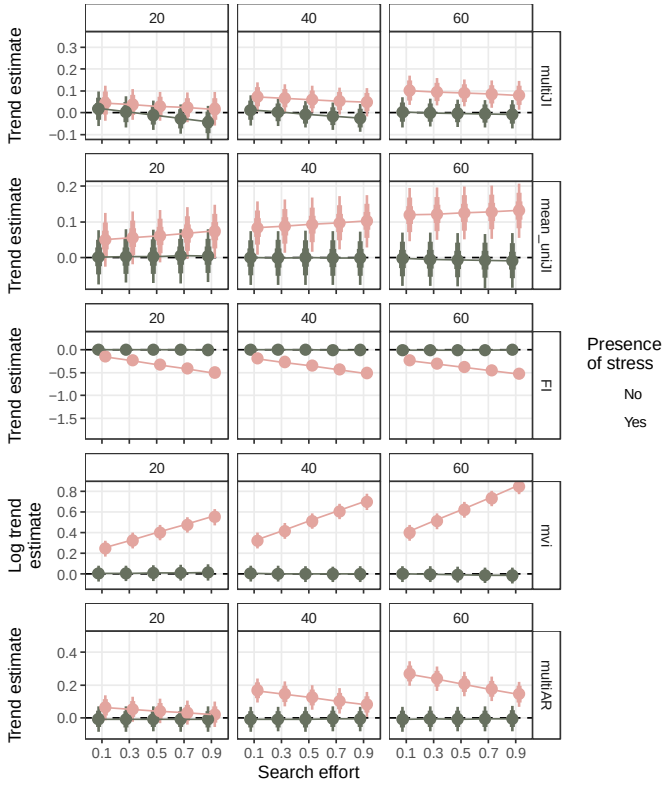


Figure 4. Posterior estimated slopes/trends of each stability metric (rows) through time, across truncated time series lengths (columns) and search efforts in 25 species communities. Dark coloured estimates represent trends in unstressed communities, while light coloured estimates are those in stressed communities. The reported values are the posterior density median values (circles), with 80% (thickest bars) and 95% (thinnest bars) credible intervals.

Threshold transgression

For the Jacobian estimating metrics multiJI, max_uniJI and multiAR, the size of the community and time series length were the best predictors of reported instability probability (Figure 5, Tables S16-S18) – i.e. displaying a metric value greater than its instability threshold. For multiJI and max_uniJI that threshold was 1 and 0 for multiAR. In all metrics,

for a constant time series length and search effort, the 25 species community always displayed a higher probability of reported instability than 15 or 5 (Figure 5) while shorter time series also increased the probability of reported instability. max_uniJI always signalled instability regardless of data quality and presence of stress, while multiJI became less likely to report false positives as data quality increased.

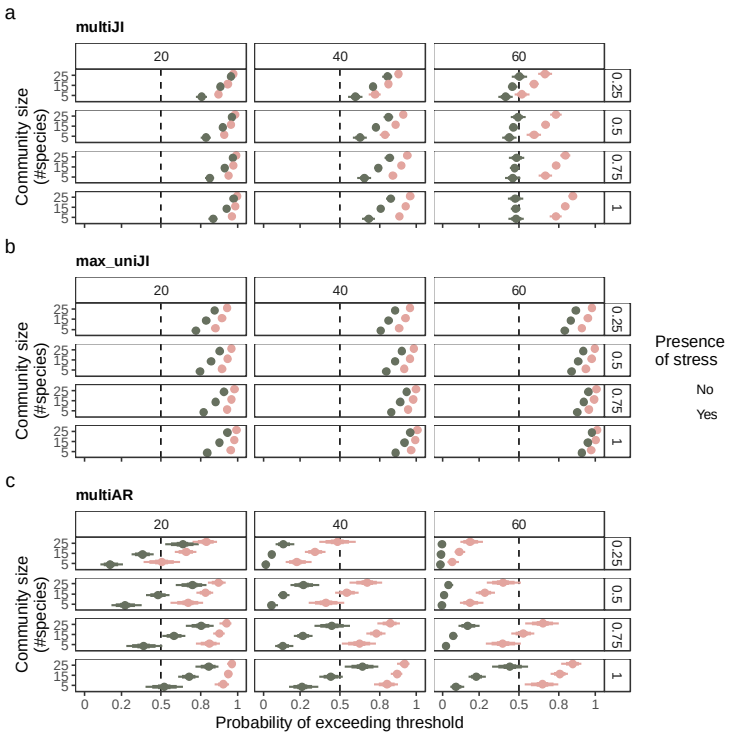


Figure 5. Posterior estimated probabilities of the stability metrics multiJI, max_uniJI and multiAR transgressing its threshold a. Estimates have been delineated by truncated time series lengths (columns) and search efforts (rows) in the three different community sizes. Dark coloured estimates represent trends in unstressed communities, while light coloured estimates are those in stressed communities. The reported values are the posterior density median values (circles), with 80% (thickest bars) and 95% (thinnest bars) credible intervals back transformed from log odds to probabilities.

Search effort did not influence the probability of instability except in multiJI predictions. Time series length generally decreased probability of reported instability but synergised with search effort so that high search effort resulted in equivalent probabilities across time series length.

If the three metrics are compared, multiJI and multiAR displayed identical behaviour in longer time series, implying both estimate similar Jacobians. Overall, multiJI displayed the highest true positive rate (highest probabilities in stressed communities) whereas max_uniJI displayed the lowest false positive rate (lowest probabilities in unstressed communities). multiJI generally balanced the ratio better than max_uniJI and multiAR with the latter most sensitive to low search efforts. multiJI was also most robust at the highest data qualities (long time series and high search efforts) across community sizes, but more conservative than the other metrics at low data qualities.

Discussion

Quantifying population and community stability/resilience loss has emerged as a crucial target of ecological research (Kéfi et al., 2019), mapping directly on to the needs of practitioners and policy makers (Donohue et al., 2016; Nicholson et al., 2012). Despite the complexity of ecological systems, political decisions often require simplifications to allow efficient decision making and action (Donatti et al., 2020; Nicholson et al., 2012). Stability metrics represent a bridge between the need for indicators of ecosystem variability, recovery, and simple user interpretation. We show here that stability metrics can perform as anticipated in simulated data analogous to real world biodiversity datasets though the choice of metric and contributing time series requires caution.

Influence of data quality and community size

Our simulations suggest that time series length has a significant impact on their ability to reconstruct estimate stability dimensions. This is anticipated behaviour (Figure 1) though the multivariate Jacobian metrics (multiJI and multiAR) unexpectedly decreased in reliability as

search effort increased. It is unclear why these metrics would behave in this way, though it is possible that, as multiJI and multiAR displayed the highest estimated autocorrelation across metrics (autocorrelation ≈ 1; Tables S1, S5, S6, S10, S11 and S15), the error introduced by low search effort acted as a low pass filter which weakened the influence of autocorrelation and allowed the true interactions between species to be identified. Similarly, our assumption that search effort linearly influences metric ability may also be violated by the strong non-linearities we observe in the raw indicator trends (Figures S1-S3).

The increasing trend in the metrics with time series length follows the typical behaviour of biodiversity indicators (Clements et al., 2015; Gonzalez et al., 2016; Lenton et al., 2012), with longer time series displaying more reliable estimates of the true trend (Wauchope et al., 2019). This effect is compounded by the caveats of Sugihara et al. (2012) that a minimum of 30 time points are required for reliable empirical dynamic modelling estimates. This is obviously unfeasible for many yearly abundance accounts reported in the LPI and other datasets, and while we show that correct estimates can be made with these short time series, their accuracy is weaker than in time series longer than 30 time points.

It is encouraging, however, that unstressed communities consistently showed no trend in any metric (Figures 2-4) although often signalled instability for metrics with a threshold (Figure 5). This implies that if a trend is observed, then that stability dimension is decreasing prior to abundance decline, and intervention is likely required. Note that no mechanism or driver of stability loss is provided by these metrics and only phenomenological signals are returned. These indicators consequently require system specific knowledge for successful intervention, but this is a recurring challenge for current abundance-based biodiversity indicators such as the LPI (Ledger et al., 2023).

Overall, the dominant determinant of metric performance was the number of species in the simulated community. The multivariate metrics estimating the communities' Jacobian (multiJI and multiAR) were particularly influenced, with weakening performance as community size increased. Ushio et al. (2018) acknowledge the potential for this degradation and suggest it is dependent on the choice of species included in the Jacobian estimation. We show here that, even when all species contributing to the system are known and measured, including all species compromises the indices' ability. We therefore support Ushio et al. (2018) that contributing species must be carefully chosen either based upon their trophic and functional importance to prevent the masking of true system behaviour. The univariate Jacobian metric (mean_uniJI) also surprisingly outperformed the multivariate (multiJI and multiAR), though it has been suggested that representative variables can provide sufficient dynamics without masking the signal with other uninformative time series (Grziwotz et al., 2023). This does however partially conflict with Donohue et al. (2013, 2016) and Kéfi et al. (2019) who clearly highlight that both uni- and multidimensional measurements of stability are required. To clarify, we agree that biodiversity and indicators need to be multidimensional, but caution against naïve species selection. Appropriate species may be selected based upon their dominance in the ecosystem in terms of abundance (Ushio et al., 2018), or inter-species interaction strengths (Medeiros et al., 2022; Medeiros & Saavedra, 2023; White et al., 2020).

A shift towards multivariate data

A general restriction of the LPI specifically is its primarily single species focus (Ledger et al., 2023; Loh et al., 2005). BIOTIME (Dornelas et al., 2018) and Global Population Dynamics Database (Inchausti & Halley, 2001) have greater assemblage contributions. That said, it is

clear that biodiversity monitoring requires multidimensional representation going forward (Donohue et al., 2013; Kéfi et al., 2019; Medeiros et al., 2022; Medeiros & Saavedra, 2023; White et al., 2020) to provide the reliable estimates required for effective governance. Such data will become more readily available with the advent of autonomous and remote ecological monitoring (Besson et al., 2022), but there will be a need for tools capable of synthesising that data in to an interpretable form for decision makers. We suggest these stability metrics can contribute to this synthesis in extension of traditional approaches such as species diversity or richness (Inchausti & Halley, 2001; Turney & Buddle, 2016) due to their development specifically focussing on species interactions. We go further, to suggest that multivariate metrics also have potential to characterise the 'trait' state-space for focal populations such a target fisheries or forestry. Using trait measurements across plots or through time, stability of a population/assemblage in terms of its traits is possible which allows ecosystem managers to ask questions regarding 'timelines of collapse' (Cerini et al., 2023) or population intactness for the individual/population/community of management interest. However, if multivariate data is not available, this study supports Grziwotz et al (2023)'s findings that univariate Jacobian estimates can also be reliable indicators of system dynamics in the absence of complete monitoring.

Comparison to critical transition indicators

A similar collection of indicators to the metrics discussed here have received extensive research interest in relation to predicting certain forms of regime shift and population collapse (Baruah et al., 2022; Clements et al., 2015; Dakos, 2018; Dakos et al., 2012). These 'Early Warning Signals' (EWSs) can be considered a sub-class of stability indicator (Dakos et al., 2015) as they quantify resilience when a system has two alternative stable states and the capability to transition/regime shift between them. With current debates over the

commonality of alternative stable states in ecology (Dakos et al., 2015; Davidson et al., 2023) and inconsistent ability in real world data (Burthe et al., 2016; Zhang, 2020), the benefit of the explored metrics over EWSs is their ability to quantify stability regardless of the presence of multiple stable states (Lyapunov, 1992; Medeiros et al., 2022; Sugihara et al., 2012). With the complex intrinsic and extrinsic interactions driving ecosystem dynamics possibly masking true multiple stability (Hillebrand et al., 2023) or ecosystems themselves experiencing long transients rather than regime shifts (Hastings et al., 2018), the Jacobian metrics discussed here are conceptually capable of accurate estimation under both heuristics. Such capability therefore expedites the general suggestions to consider non-equilibrium resilience and stability change in ecological research (Donohue et al., 2013, 2016; Kéfi et al., 2019).

Trends vs thresholds

A key challenge for designing indicators appropriate for policy is that governance primarily focusses on thresholds (Donohue et al., 2016; United Nations, 2020) as they allow unambiguous interpretation and enforcement. Conversely, trends require sufficient subjectivity that the vague language associated with trend reporting results in broad objective language (United Nations, 2020). Many governments therefore use indicators with set target thresholds for biodiversity rather than considering a continuous gradient of ‘state’ which is suggested to be more appropriate for modern coupled socio-ecological systems (Hillebrand et al., 2023).

This requirement for a threshold is also explored by the LPI. The debates surrounding the LPI and its interpretation (Dornelas et al., 2019; Leung et al., 2020) ultimately revolve around reconciling the variability of local trend directions (winner and loser populations) with the

short time series typically available (Gonzalez et al., 2016). Indeed, controlling for co-dependencies between separate time series such as temporal, spatial and phylogenetic autocorrelation increases the uncertainty of trend estimates (Johnson et al., 2024). This indicator/monitoring uncertainty erodes legislators' ability to make confident governance, interacts with other forms of uncertainty (e.g. political strategies, lobbying effort, pre-existing laws (Dewulf & Biesbroek, 2018)) and leads to broad objectives. The Jacobian estimating metrics we explored minimise this incongruency between trends and thresholds by facilitating both forms of interpretation – an increase in trend is indicative of decreasing stability, while exceeding of the threshold value represents detection of instability. Meanwhile, FI and mvi do not express thresholds but their superior ability relative to Jacobian indices in this study implies trends are more reliable monitoring instruments for ecosystem stability than instability thresholds.

Indeed, instability is not necessarily atypical of natural systems. Lake fish communities can display consistent periods of instability though the year (Ushio et al., 2018), Canadian boreal forests are temporarily unstable following historic fires (Héon et al., 2014), and experimental intertidal communities recover (i.e. display transient instability) following single pulse stress (White et al., 2020). In the metrics tested here, we identified that instability is regularly reported for shorter time series – presumably resulting from transient dynamics within the data window - and that a longer period is necessary for the stable systems to be correctedly classified as such. Stability metric thresholds are therefore only informative if instability is prolonged.

Conclusion

Our study demonstrates stability metrics can accurately quantify stability of simulated multispecies communities manipulated into realistic data qualities. Contributing species/time series is the key consideration for these metrics' functionality and requires ecological knowledge input from the user. Stability is strongly suggested to be under characterised during global biodiversity monitoring (Capdevila et al., 2022; Kéfi et al., 2019) and the metrics tested here have the potential to contribute generic inference from abundance data alone. We consequently anticipate Jacobian estimation techniques to increase in popularity but pre-empt their use in biodiversity datasets by cautioning towards the longest time series possible and ecologically important species.

References

Ahmad, N., Derrible, S., Eason, T., & Cabezas, H. (2016). Using Fisher Information to Track Stability in Multivariate Systems. *Royal Society Open Science*, *3*(11), 160582. https://doi.org/10.1098/rsos.160582

Baruah, G., Ozgul, A., & Clements, C. F. (2022). Community Structure Determines the Predictability of Population Collapse. *Journal of Animal Ecology*, *n/a*(n/a). https://doi.org/10.1111/1365-2656.13769

Besson, M., Alison, J., Bjerge, K., Gorochowski, T. E., Høye, T. T., Jucker, T., Mann, H. M. R., & Clements, C. F. (2022). Towards the fully automated monitoring of ecological communities. *Ecology Letters*, *25*(12), 2753–2775. https://doi.org/10.1111/ele.14123

Brock, W., & Carpenter, S. R. (2006). Variance as a Leading Indicator of Regime Shift in Ecosystem Services. *Ecology and Society*, *11*, 9.

Brotons, L., Pou, N., Sainz de la Maza, P., & Pont, S. (2020). *Estat de la Natura a Catalunya 2020*. Departament de Territori i Sostenibilitat, Generalitat de Catalunya.

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. (2016). Do early warning indicators consistently predict nonlinear change in long-term ecological data? *Journal of Applied Ecology*, *53*(3), 666–676. https://doi.org/10.1111/1365-2664.12519

Capdevila, P., Noviello, N., McRae, L., Freeman, R., & Clements, C. F. (2022). Global patterns of resilience decline in vertebrate populations. *Ecology Letters*, *25*(1), 240–251. https://doi.org/10.1111/ele.13927

Cardinale, B. J., Duffy, J. E., Gonzalez, A., Hooper, D. U., Perrings, C., Venail, P., Narwani, A., Mace, G. M., Tilman, D., Wardle, D. A., Kinzig, A. P., Daily, G. C., Loreau, M., Grace, J. B., Larigauderie, A., Srivastava, D. S., & Naeem, S. (2012). Biodiversity loss and its impact on humanity. *Nature*, *486*(7401), 59–67. https://doi.org/10.1038/nature11148

Ceballos, G., Ehrlich, P. R., Barnosky, A. D., García, A., Pringle, R. M., & Palmer, T. M. (2015). Accelerated modern human–induced species losses: Entering the sixth mass extinction. *Science Advances*, *1*(5), e1400253. https://doi.org/10.1126/sciadv.1400253

Cerini, F., Childs, D. Z., & Clements, C. F. (2023). A predictive timeline of wildlife population collapse. *Nature Ecology & Evolution*, *7*(3), 320–331. https://doi.org/10.1038/s41559-023-01985-2

Chang, C.-W., Ushio, M., & Hsieh, C. (2017). Empirical Dynamic Modeling for Beginners. *Ecological Research*, *32*(6), 785–796. https://doi.org/10.1007/s11284-017-1469-9

Christopoulos, D. (2016). On the Efficient Identification of an Inflection Point. *INTERNATIONAL JOURNAL OF MATHEMATICS AND SCIENTIFIC COMPUTING (ISSN: 2231-5330)*, *6*.

Clements, C. F., Drake, J. M., Griffiths, J. I., & Ozgul, A. (2015). Factors Influencing the Detectability of Early Warning Signals of Population Collapse. *The American Naturalist*, *186*(1), 50–58. https://doi.org/10.1086/681573

Dakos, V. (2018). Identifying Best-Indicator Species for Abrupt Transitions in Multispecies Communities. *Ecological Indicators*, *94*, 494–502. https://doi.org/10.1016/j.ecolind.2017.10.024

Dakos, V., Carpenter, S. R., Brock, W. A., Ellison, A. M., Guttal, V., Ives, A. R., Kéfi, 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*(7), e41010. https://doi.org/10.1371/journal.pone.0041010

Dakos, V., Carpenter, S. R., van Nes, E. H., & Scheffer, M. (2015). Resilience indicators: Prospects and limitations for early warnings of regime shifts. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1659), 20130263. https://doi.org/10.1098/rstb.2013.0263

Dale, V. H., & Beyeler, S. C. (2001). Challenges in the development and use of ecological indicators. *Ecological Indicators*, *1*(1), 3–10. https://doi.org/10.1016/S1470-160X(01)00003-6

Davidson, T. A., Bennion, H., Reid, M., Sayer, C. D., & Whitmore, T. J. (2018). Towards better integration of ecology in palaeoecology: From proxies to indicators, from inference to understanding. *Journal of Paleolimnology*, *60*(2), 109–116. https://doi.org/10.1007/s10933-018-0032-1

Davidson, T. A., Sayer, C. D., Jeppesen, E., Søndergaard, M., Lauridsen, T. L., Johansson, L. S., Baker, A., & Graeber, D. (2023). Bimodality and alternative equilibria do not help

explain long-term patterns in shallow lake chlorophyll-a. *Nature Communications*, *14*(1), 398. https://doi.org/10.1038/s41467-023-36043-9

Dewulf, A., & Biesbroek, R. (2018). Nine lives of uncertainty in decision-making: Strategies for dealing with uncertainty in environmental governance. *Policy and Society*, *37*(4), 441–458. https://doi.org/10.1080/14494035.2018.1504484

Doherty, T. S., Glen, A. S., Nimmo, D. G., Ritchie, E. G., & Dickman, C. R. (2016). Invasive predators and global biodiversity loss. *Proceedings of the National Academy of Sciences*, *113*(40), 11261–11265. https://doi.org/10.1073/pnas.1602480113

Donatti, C. I., Harvey, C. A., Hole, D., Panfil, S. N., & Schurman, H. (2020). Indicators to measure the climate change adaptation outcomes of ecosystem-based adaptation. *Climatic Change*, *158*(3), 413–433. https://doi.org/10.1007/s10584-019-02565-9

Donohue, I., Hillebrand, H., Montoya, J. M., Petchey, O. L., Pimm, S. L., Fowler, M. S., Healy, K., Jackson, A. L., Lurgi, M., McClean, D., O'Connor, N. E., O'Gorman, E. J., & Yang, Q. (2016). Navigating the complexity of ecological stability. *Ecology Letters*, *19*(9), 1172–1185. https://doi.org/10.1111/ele.12648

Donohue, I., Petchey, O. L., Montoya, J. M., Jackson, A. L., McNally, L., Viana, M., Healy, K., Lurgi, M., O'Connor, N. E., & Emmerson, M. C. (2013). On the dimensionality of ecological stability. *Ecology Letters*, *16*(4), 421–429. https://doi.org/10.1111/ele.12086

Dornelas, M., Antão, L. H., Moyes, F., Bates, A. E., Magurran, A. E., Adam, D., Akhmetzhanova, A. A., Appeltans, W., Arcos, J. M., Arnold, H., Ayyappan, N., Badihi, G., Baird, A. H., Barbosa, M., Barreto, T. E., Bässler, C., Bellgrove, A., Belmaker, J., Benedetti-Cecchi, L., … Zettler, M. L. (2018). BioTIME: A database of biodiversity time series for the Anthropocene. *Global Ecology and Biogeography*, *27*(7), 760–786. https://doi.org/10.1111/geb.12729

Dornelas, M., Gotelli, N. J., Shimadzu, H., Moyes, F., Magurran, A. E., & McGill, B. J. (2019). A balance of winners and losers in the Anthropocene. *Ecology Letters*, *22*(5), 847–854. https://doi.org/10.1111/ele.13242

Dutta, P. S., Sharma, Y., & Abbott, K. C. (2018). Robustness of Early Warning Signals for Catastrophic and Non-Catastrophic Transitions. *Oikos*, *127*(9), 1251–1263. https://doi.org/10.1111/oik.05172

Environment and Climate Change Canada. (2019). *Canadian Environmental Sustainability Indicators: Canadian species index*.

Gonzalez, A., Cardinale, B. J., Allington, G. R. H., Byrnes, J., Arthur Endsley, K., Brown, D. G., Hooper, D. U., Isbell, F., O'Connor, M. I., & Loreau, M. (2016). Estimating local biodiversity change: A critique of papers claiming no net loss of local diversity. *Ecology*, *97*(8), 1949–1960. https://doi.org/10.1890/15-1759.1

Grziwotz, F., Chang, C.-W., Dakos, V., van Nes, E. H., Schwarzländer, M., Kamps, O., Heßler, M., Tokuda, I. T., Telschow, A., & Hsieh, C. (2023). Anticipating the Occurrence and Type of Critical Transitions. *Science Advances*, *9*(1), eabq4558. https://doi.org/10.1126/sciadv.abq4558

Hasselmann, K. (1988). PIPs and POPs: The reduction of complex dynamical systems using principal interaction and oscillation patterns. *Journal of Geophysical Research: Atmospheres*, *93*(D9), 11015–11021. https://doi.org/10.1029/JD093iD09p11015

Hastings, A. (2016). Timescales and the management of ecological systems. *Proceedings of the National Academy of Sciences*, *113*(51), 14568. https://doi.org/10.1073/pnas.1604974113

Hastings, A., Abbott, K. C., Cuddington, K., Francis, T., Gellner, G., Lai, Y.-C., Morozov, A., Petrovskii, S., Scranton, K., & Zeeman, M. L. (2018). Transient phenomena in ecology. *Science*, *361*(6406), eaat6412. https://doi.org/10.1126/science.aat6412

Héon, J., Arseneault, D., & Parisien, M.-A. (2014). Resistance of the boreal forest to high burn rates. *Proceedings of the National Academy of Sciences*, *111*(38), 13888–13893. https://doi.org/10.1073/pnas.1409316111

Hillebrand, H., Kuczynski, L., Kunze, C., Rillo, M. C., & Dajka, J.-C. (2023). Thresholds and tipping points are tempting but not necessarily suitable concepts to address anthropogenic biodiversity change—An intervention. *Marine Biodiversity*, *53*(3), 43. https://doi.org/10.1007/s12526-023-01342-3

Hofbauer, J., & Sigmund, K. (1998). *Evolutionary Games and Population Dynamics*. Cambridge University Press. https://doi.org/10.1017/CBO9781139173179

Holling, C. S. (1973). Resilience and stability of ecological systems. *Annual Review of Ecology and Systematics*, *4*(1), 1–23. https://doi.org/10.1146/annurev.es.04.110173.000245

Hughes, B. B., Beas-Luna, R., Barner, A. K., Brewitt, K., Brumbaugh, D. R., Cerny-Chipman, E. B., Close, S. L., Coblentz, K. E., de Nesnera, K. L., Drobnitch, S. T., Figurski, J. D., Focht, B., Friedman, M., Freiwald, J., Heady, K. K., Heady, W. N., Hettinger, A., Johnson, A., Karr, K. A., … Carr, M. H. (2017). Long-Term Studies Contribute Disproportionately to Ecology and Policy. *BioScience*, *67*(3), 271–281. https://doi.org/10.1093/biosci/biw185

Inchausti, P., & Halley, J. (2001). Investigating long-term ecological variability using the Global Population Dynamics Database. *Science*, *293*(5530), 655–657. https://doi.org/10.1126/science.293.5530.655

IUCN. (2022). *The IUCN Red List of Threatened Species* (Versions 2022-2) [Dataset]. https://www.iucnredlist.org

Johnson, T. F., Beckerman, A. P., Childs, D. Z., Webb, T. J., Evans, K. L., Griffiths, C. A., Capdevila, P., Clements, C. F., Besson, M., Gregory, R. D., Thomas, G. H., Delmas,

E., & Freckleton, R. P. (2024). Revealing uncertainty in the status of biodiversity change. *Nature*, *628*(8009), 788–794. https://doi.org/10.1038/s41586-024-07236-z

Karunanithi, A. T., Cabezas, H., Frieden, B. R., & Pawlowski, C. W. (2008). Detection and Assessment of Ecosystem Regime Shifts from Fisher Information. *Ecology and Society*, *13*(1). http://www.jstor.org/stable/26267924

Kéfi, S., Domínguez-García, V., Donohue, I., Fontaine, C., Thébault, E., & Dakos, V. (2019). Advancing our understanding of ecological stability. *Ecology Letters*, *22*(9), 1349–1356. https://doi.org/10.1111/ele.13340

Ledger, S. E. H., Loh, J., Almond, R., Böhm, M., Clements, C. F., Currie, J., Deinet, S., Galewski, T., Grooten, M., Jenkins, M., Marconi, V., Painter, B., Scott-Gatty, K., Young, L., Hoffmann, M., Freeman, R., & McRae, L. (2023). Past, present, and future of the Living Planet Index. *Npj Biodiversity*, *2*(1), 12. https://doi.org/10.1038/s44185-023-00017-3

Lenton, T. M., Livina, V. N., Dakos, V., van Nes, E. H., & Scheffer, M. (2012). Early Warning of Climate Tipping Points from Critical Slowing down: Comparing Methods to Improve Robustness. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, *370*(1962), 1185–1204. https://doi.org/10.1098/rsta.2011.0304

Leung, B., Hargreaves, A. L., Greenberg, D. A., McGill, B., Dornelas, M., & Freeman, R. (2020). Clustered versus catastrophic global vertebrate declines. *Nature*, *588*(7837), 267–271. https://doi.org/10.1038/s41586-020-2920-6

Loh, J., Green, R. E., Ricketts, T., Lamoreux, J., Jenkins, M., Kapos, V., & Randers, J. (2005). The Living Planet Index: Using species population time series to track trends in biodiversity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1454), 289–295. https://doi.org/10.1098/rstb.2004.1584

Loreau, M., Cardinale, B. J., Isbell, F., Newbold, T., O'Connor, M. I., & de Mazancourt, C. (2022). Do not downplay biodiversity loss. *Nature*, *601*(7894), E27–E28. https://doi.org/10.1038/s41586-021-04179-7

Lyapunov, A. M. (1992). The general problem of the stability of motion. *International Journal of Control*, *55*(3), 531–534. https://doi.org/10.1080/00207179208934253

McRae, L., Deinet, S., & Freeman, R. (2017). The Diversity-Weighted Living Planet Index: Controlling for Taxonomic Bias in a Global Biodiversity Indicator. *PLOS ONE*, *12*(1), e0169156. https://doi.org/10.1371/journal.pone.0169156

Medeiros, L. P., Allesina, S., Dakos, V., Sugihara, G., & Saavedra, S. (2022). Ranking Species Based on Sensitivity to Perturbations under Non-Equilibrium Community Dynamics. *Ecology Letters*, *n/a*(n/a). https://doi.org/10.1111/ele.14131

Medeiros, L. P., & Saavedra, S. (2023). Understanding the state-dependent impact of species correlated responses on community sensitivity to perturbations. *Ecology*, *n/a*(n/a), e4115. https://doi.org/10.1002/ecy.4115

Nicholson, E., Collen, B., Barausse, A., Blanchard, J. L., Costelloe, B. T., Sullivan, K. M. E., Underwood, F. M., Burn, R. W., Fritz, S., Jones, J. P. G., McRae, L., Possingham, H. P., & Milner-Gulland, E. J. (2012). Making Robust Policy Decisions Using Global Biodiversity Indicators. *PLOS ONE*, *7*(7), e41128. https://doi.org/10.1371/journal.pone.0041128

O'Brien, D. A., Deb, S., Sidheekh, S., Krishnan, N. C., Sharathi Dutta, P., & Clements, C. F. (2023). EWSmethods: An R package to forecast tipping points at the community level using early warning signals, resilience measures, and machine learning models. *Ecography*, *n/a*(n/a), e06674. https://doi.org/10.1111/ecog.06674

Oliver, T. H., Isaac, N. J. B., August, T. A., Woodcock, B. A., Roy, D. B., & Bullock, J. M. (2015). Declining resilience of ecosystem functions under biodiversity loss. *Nature Communications*, *6*(1), 10122. https://doi.org/10.1038/ncomms10122

O'Loughlin, L. S., & Green, P. T. (2017). Secondary invasion: When invasion success is contingent on other invaders altering the properties of recipient ecosystems. *Ecology and Evolution*, *7*(19), 7628–7637. https://doi.org/10.1002/ece3.3315

Park, J., Smith, C., Sugihara, G., & Deyle, E. (2021). *rEDM: Empirical Dynamic Modeling ('EDM')* [Computer software]. https://CRAN.R-project.org/package=rEDM

Pimm, S. L. (1984). The complexity and stability of ecosystems. *Nature*, *307*(5949), 321–326. https://doi.org/10.1038/307321a0

R Core Team. (2022). *R: A Language and Environment for Statistical Computing* [Computer software]. R Foundation for statistical computing.

Rue, H., Martino, S., & Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *71*(2), 319–392. https://doi.org/10.1111/j.1467-9868.2008.00700.x

Strogatz, S. H. (2015). *Nonlinear dynamics and chaos: With applications to physics, biology, chemistry, and engineering* (2nd ed.). CRC Press. https://doi.org/10.1201/9780429492563

Sugihara, G., May, R., Ye, H., Hsieh, C., Deyle, E., Fogarty, M., & Munch, S. (2012). Detecting Causality in Complex Ecosystems. *Science*, *338*(6106), 496 LP – 500. https://doi.org/10.1126/science.1227079

Traill, L. W., Perhans, K., Lovelock, C. E., Prohaska, A., McFallan, S., Rhodes, J. R., & Wilson, K. A. (2011). Managing for change: Wetland transitions under sea-level rise

and outcomes for threatened species. *Diversity and Distributions*, *17*(6), 1225–1233. https://doi.org/10.1111/j.1472-4642.2011.00807.x

Turney, S., & Buddle, C. M. (2016). Pyramids of species richness: The determinants and distribution of species diversity across trophic levels. *Oikos*, *125*(9), 1224–1232. https://doi.org/10.1111/oik.03404

Tylianakis, J. M., Didham, R. K., Bascompte, J., & Wardle, D. A. (2008). Global change and species interactions in terrestrial ecosystems. *Ecology Letters*, *11*(12), 1351–1363. https://doi.org/10.1111/j.1461-0248.2008.01250.x

United Nations. (1992, June 5). *Convention on biological diversity*.

United Nations. (2020). *Aichi Biodiversity Targets*. https://www.cbd.int/sp/targets/

Ushio, M., Hsieh, C., Masuda, R., Deyle, E. R., Ye, H., Chang, C.-W., Sugihara, G., & Kondoh, M. (2018). Fluctuating Interaction Network and Time-Varying Stability of a Natural Fish Community. *Nature*, *554*(7692), 360–363. https://doi.org/10.1038/nature25504

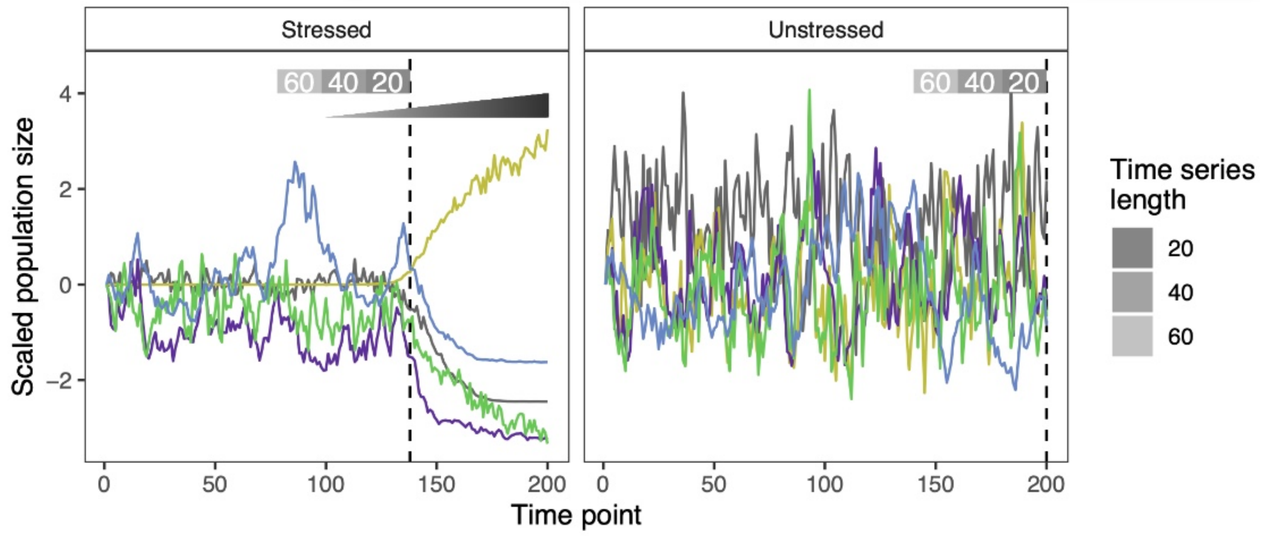
Wauchope, H. S., Amano, T., Sutherland, W. J., & Johnston, A. (2019). When can we trust population trends? A method for quantifying the effects of sampling interval and duration. *Methods in Ecology and Evolution*, *10*(12), 2067–2078. https://doi.org/10.1111/2041-210X.13302

White, L., O'Connor, N. E., Yang, Q., Emmerson, M. C., & Donohue, I. (2020). Individual species provide multifaceted contributions to the stability of ecosystems. *Nature Ecology & Evolution*, *4*(12), 1594–1601. https://doi.org/10.1038/s41559-020-01315-w

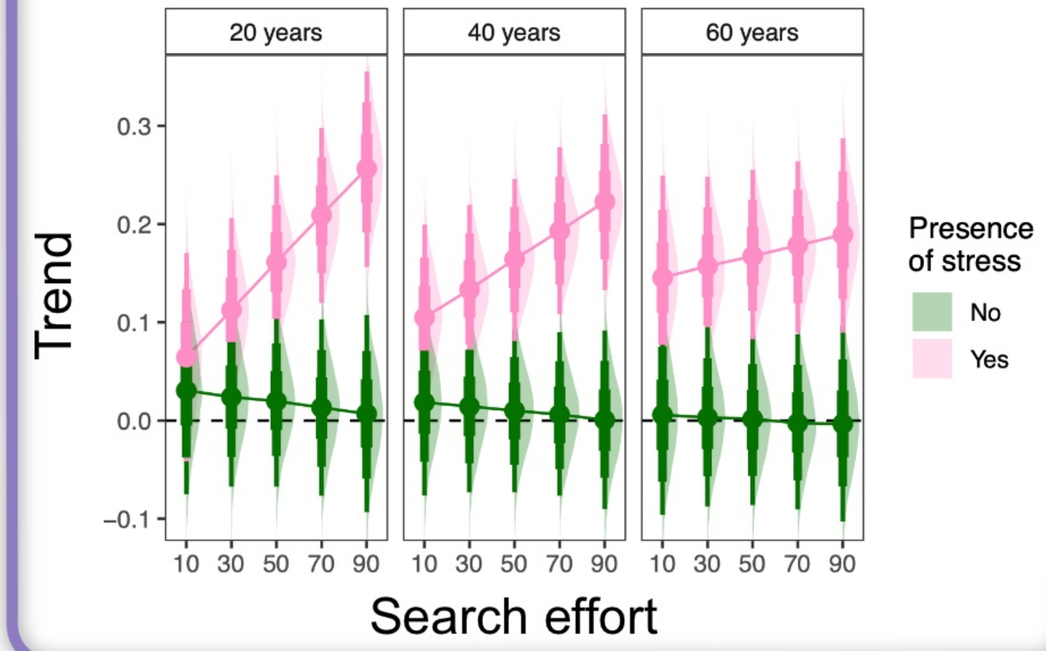
Williamson, M. S., & Lenton, T. M. (2015). Detection of bifurcations in noisy coupled systems from multiple time series. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *25*(3), 036407. https://doi.org/10.1063/1.4908603

Zhang, F. (2020). Early warning signals of population productivity regime shifts in global fisheries. *Ecological Indicators*, *115*, 106371. https://doi.org/10.1016/j.ecolind.2020.106371

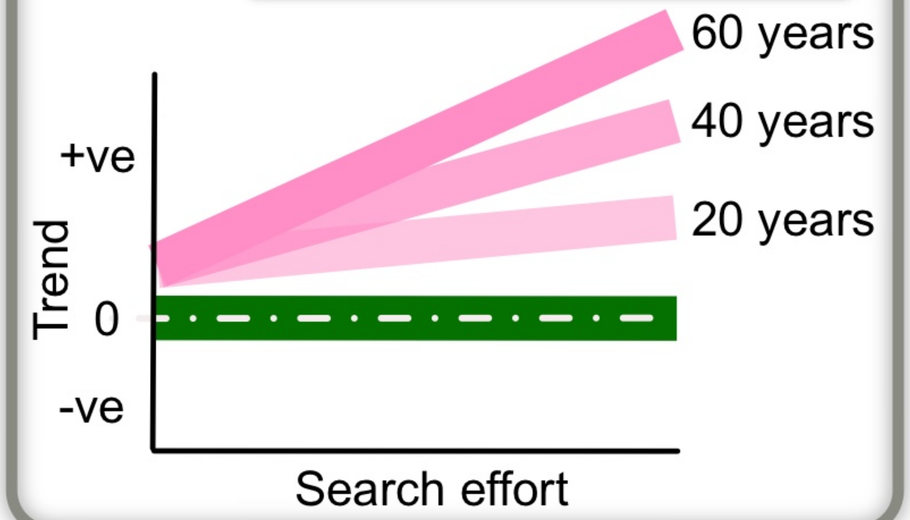
1a) In-silico experiment



2) Outcome



1b) Expectation



3) Practitioner recommendations

Stability metric trends are robust for general inference with thresholds prone to false positive errors

Longer time series are desirable as they improve reliability

Select species using ecological knowledge to mitigate against large community sizes