nature climate change

Perspective

https://doi.org/10.1038/s41558-025-02328-8

Ambiguity of early warning signals for climate tipping points

Received: 4 July 2023

Accepted: 31 March 2025

Published online: 29 April 2025

Check for updates

Max Rietkerk [1,4] ✉, Vanessa Skiba [2,4] ✉, Els Weinans [1], Raphaël Hébert [2] & Thomas Laepple [2,3]

There is concern that climate change might lead to abrupt and irreversible changes in parts of the Earth system at so-called tipping points. Theoretical considerations suggest that statistical measures can be used to detect early warning signals (EWSs) for reduced resilience, which could be interpreted as an increased proximity to climate tipping points. Here we discuss limitations of commonly used EWSs and their detection and discuss how alternative explanations can lead to resilience loss in the absence of tipping points. We argue for better testing of the existence of tipping points, beyond the application of EWSs, and propose a method to better quantify the probability of approaching tipping points using EWSs.

On the basis of palaeoclimate and model-based studies, many elements of the Earth system, such as the Greenland Ice Sheet, the Amazon rainforest, the Atlantic Meridional Overturning Circulation (AMOC) and the West Antarctic Ice Sheet, are considered prone to abrupt and irreversible changes at so-called tipping points[1–3]. Crossing tipping points may lead to high-impact and catastrophic outcomes[4], involving such changes of climate system elements. However, quantifying the probability of tipping is challenging, and the latest IPCC assessment report assigned low to medium confidence for most of these climate tipping points to be reached in the twenty-first century, with their likelihood of occurrence generally increasing at higher warming levels[5].

Dynamical systems approaching a tipping point recover more slowly from small perturbations[6] (Box 1). This phenomenon is termed 'critical slowing down' and is associated with increasing memory and variability of the system, which are interpreted as early warning signals (EWSs) for tipping points. These can be detected via changes in statistical properties of time series of variables from the system under study, and various metrics for that have been developed (Box 1). Recently, such EWS metrics have been used to infer approaching tipping points in Earth system components (for example, refs. 7–10). However, detecting and interpreting EWSs in empirical data comes with multiple conceptual and statistical issues (Table 1), challenging the interpretation of the results (for example, refs. 11–13).

In this Perspective, we reflect on these uncertainties regarding EWSs and associated metrics to detect approaching climate tipping points in Earth system elements from observational data. We review the confidence regarding the existence of climate tipping points and the lack of mechanistic understanding of EWSs. Furthermore, we discuss the potential misinterpretation of EWSs due to overlooked alternative explanations and the pitfalls related to theoretical and statistical challenges (Table 1) that might occur when applying EWSs to observational data. To address these challenges, we propose a statistical framework for the application of EWSs and provide recommendations for future efforts.

We argue that it is crucial to enhance the robustness in the application of EWS theory to reduce the probability of false alarms. While we critically discuss the ability of EWSs to anticipate tipping points, this should not be misinterpreted as diminishing the urgency of climate mitigation. Given the potential for catastrophic outcomes, the precautionary principle necessitates proactive measures even in the face of scientific uncertainty.

Tipping points and positive feedback

Tipping points are critical thresholds where a dynamical system with alternative stable states abruptly shifts from one stable state to the other, due to a relatively small change in environmental conditions[1,14] (Box 1). Because of the causal relation between the occurrence of tipping points and positive feedback, the reversal of the system's state after crossing the tipping point is hindered, a phenomenon called hysteresis[1,6] (Box 1). Positive feedbacks are ubiquitous in climate system elements, such as the Greenland Ice Sheet, the Amazon rainforest

[1]Department of Environmental Sciences, Copernicus Institute of Sustainable Development, Utrecht University, Utrecht, the Netherlands. [2]Alfred-Wegener-Institut, Helmholtz-Zentrum für Polar- und Meeresforschung, Potsdam, Germany. [3]MARUM–Center for Marine Environmental Sciences and Faculty of Geosciences, University of Bremen, Bremen, Germany. [4]These authors contributed equally: Max Rietkerk, Vanessa Skiba.
✉e-mail: m.g.rietkerk@uu.nl; vanessa.skiba@awi.de

https://doi.org/10.1038/s41558-025-02328-8

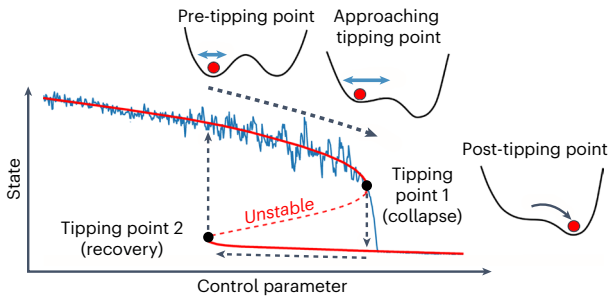
BOX 1

Tipping points and their anticipation using statistical EWSs

Tipping points (black dots; see example from the Ricker model[99] in the figure) are critical thresholds at which a small change in a control parameter (*x* axis) can lead to a shift of a system from one stable state to another. Such shifts with tipping points are synonymously called catastrophic shifts or critical transitions (as opposed to non-catastrophic shifts or non-critical transitions without tipping points)[1,6,12,100] and are indicated in the figure by the dashed black arrows starting at the tipping points and moving towards the other stable state. Tipping points are caused by positive feedbacks, and reversal of these critical transitions is associated with hysteresis. This is illustrated in the figure where the collapse to the lower system state at tipping point 1 occurs at a higher value of the control parameter compared with the recovery to the higher system state at tipping point 2. Statistical EWSs have been proposed to detect such tipping points, on the basis of the phenomenon of critical slowing down, that is, increased recovery time of system state variables (for example, refs. 6,100,101).

Critical slowing down results in an increased amplitude of fluctuations of system state variables as the system approaches a tipping point. This is indicated by the blue fluctuating line in the figure, also depicted as blue double arrows in the inset sketches. The blue line fluctuates around an equilibrium of a system, illustrated by the red line in the figure, also shown by the position of the red balls in the inset sketches. This indicates an ongoing destabilization of a system when approaching a tipping point.

EWSs can be found in a system that is subject to critical slowing down. Variables of this system will exhibit changes in their statistical properties, in particular increases in amplitude of fluctuations and memory, which can be used as EWSs (for example, ref. 63). Those EWSs can be estimated statistically via variance, lag-1 autocorrelation (AC1) or the restoring rate[6,38,50]. Various more recent metrics based on this principle have been developed (for example, refs. 8,42,47,102–104). Much recent work has focused on the detection of critical slowing down in climatic and environmental data using EWS metrics (for example, refs. 7,8,10,34,59,105). Those metrics are usually computed in a running window following detrending of the time series (the trend is illustrated as the red equilibrium line in the figure). If an increase in the metrics over time is detected, this is interpreted as EWSs indicating critical slowing down and an approaching tipping point.



and the AMOC[2]. An example of such a process, where a change leads to further increase of that change, is the evapotranspiration–rainfall feedback in the Amazon. There such positive feedback of reduced forest cover would lead to decreased evapotranspiration, causing less rainfall, leading to even further reduced forest cover.

This type of tipping point is related to fold bifurcations, which occur due to changes in stability properties of a dynamical system. Such bifurcation-induced tipping can be detected using EWSs[6] (Box 1) and is therefore the focus of this Perspective. We do not consider noise-induced and rate-induced tipping, as these types of tipping cannot be anticipated by EWSs[14,15]. Furthermore, we exclude other, qualitatively different, bifurcations with no causal relation with positive feedback from the tipping point definition used in this Perspective. Examples of such different bifurcations are smooth (non-abrupt), gradual and potentially reversible transitions, such as the approximate linear variation of Arctic summer sea ice with global surface temperature, or the emergence of regular cycles, such as the onset of climate oscillations[1,5,14].

Model dependence of tipping behaviour

To obtain a mechanistic understanding of feedbacks, such systems dynamics are often reduced to simple models, whereby positive feedbacks drive the existence of tipping points (for example, refs. 16,17). In such 'idealized' or 'minimal' models[1,16], positive feedbacks are often isolated as mechanisms to study their effects on properties of system variables, typically through varying parameter values determining the strength of feedbacks[1,17,18]. Thus, such simple models are well suited to infer causal relations between the occurrence of tipping points and the strength of positive feedbacks.

In contrast to the simplified models often used in tipping-point analysis, tipping points are less often found in more comprehensive and high-complexity Earth system models (ESMs). A systematic search for regional abrupt changes has been conducted previously in an ESM ensemble to detect potential tipping points for different global warming scenarios[19]. This revealed that 'abrupt shifts' of each identified climate system element were observed in only a minority of these models, and mostly under the worst-case high-emissions scenario, representative concentration pathway 8.5[19] (compare table S2 in ref. 3). Moreover, detected abrupt shifts are not necessarily the result of tipping points, for example, when shifts of climate system elements

Table 1 | Theoretical and statistical challenges addressed in this Perspective related to the use of EWSs for climate tipping points

Challenge	References	Theoretical	Statistical
Complex noise regimes	39,40,42,96,97	X	X
Timescales of system relative to driver	44,45	X	
Critical slowing down occurs also without tipping points	11,12,98	X	
Changes in data availability, quality and processing techniques	61,64–67		X
Spatial autocorrelation	73		X
Changes in climate variability or random events (for example, volcanic eruptions)	41,68–71	X	X
Changes in observational system (for example, bioturbation in sediment cores)	74	X	X

https://doi.org/10.1038/s41558-025-02328-8

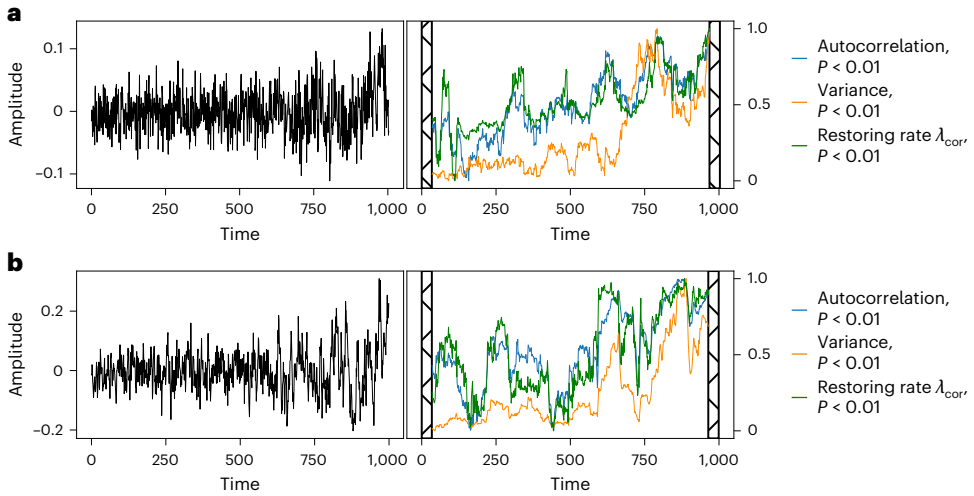


Fig. 1 | EWSs can occur with and without tipping points. EWSs can arise also in the absence of a tipping point, demonstrated here on the basis of a simple AMOC model[17]. The model describes salinity differences, dependent on freshwater flux and strength of a positive salinity–advection feedback. With increasing freshwater flux over time, for a strong positive feedback there is a critical transition with a tipping point, and for a weak positive feedback there is a smooth, non-critical transition (not abrupt and reversible) without a tipping point. The autocorrelation variance, and corrected restoring rate λ_{cor} (ref. 8) all show a significant increase and thus provide EWSs, for both the critical transition with a tipping point and the smooth transition without a tipping point. The EWSs are all scaled between 0 and 1 for ease of visualization. **a**, Time series of salinity differences for increasing freshwater flux over time towards a critical transition with a tipping point and associated EWS metrics. **b**, Time series of salinity differences for increasing freshwater flux over time towards a smooth transition without a tipping point and associated EWS metrics. See Methods for details.

are reversible (that is, without hysteresis; Box 1), such as demonstrated for Arctic sea-ice loss[20]. So, although abrupt shifts can be detected, there is so far only limited evidence from high-complexity ESMs for the existence of climate tipping points. Further model intercomparisons with improved ESMs will help to clarify this evidence and refine the assessment of the likelihood of climate tipping points[3].

As an example, tipping of the AMOC has commonly been demonstrated with simple models[16,17,21,22], but rarely in high-complexity ESMs[19,23] (compare table S2 in ref. 3, but see refs. 24,25). This is probably because increasing model complexity and spatiotemporal resolution leads to smoother transitions and more stability in the AMOC[18,26–28]. It is plausible that stabilizing mechanisms operating in high-complexity ESMs may counteract or override positive feedbacks, thereby preventing tipping points. In addition, spatial effects such as spatial heterogeneity can obscure or prevent the occurrence of tipping points[28]. For example, the coexistence of different stable states in one spatial domain is associated with multi-stability, that is, multiplicity of stable states[28]. One example is the possibility of multiple states of the AMOC, besides the convective and non-convective states that are seen in simplified models. Convection may also occur only in various parts of the spatial domain, depending on its size and spatial heterogeneity[29]. This may lead to fragmented or intermediate tipping, or partial spatial reorganization of the system[18,29]. Therefore, we suggest that counteracting interactions at various spatiotemporal scales could explain the lack of clear evidence for tipping points in climate system elements in high-complexity ESMs. However, ESMs also have structural uncertainties and could be regionally too stable (for example, ref. 30).

Various studies based on conceptual models find that stabilizing and destabilizing interactions between potential tipping elements, for example, the Greenland Ice Sheet, the AMOC and the Amazon rainforest, could increase the probability of tipping and invoke tipping cascades[4,31–35]. This suggests that destabilizing interactions between climate system elements could dominate over stabilizing ones[4,36,37]. However, studies investigating climate system element interactions so far are based only on simple conceptual models. We postulate that the intersystem and large-scale emerging stability of complex systems addressed here might not be fully captured by such models. Thus, it is still unknown how climate system elements interact in complex ESMs and in the real global Earth system.

Theoretical challenges for detecting EWSs

To formulate dynamical models, assumptions have to be made about which processes to implement and which to ignore. Under these assumptions, the dynamics of a system may be reflected by a simple model, such as illustrated in Box 1. We know that in such simple systems, and with relatively simple noise regimes, slowing is a relevant process that can be detected by an increase in metrics for autocorrelation and variability[6] (Box 1), as well as by measures more directly related to slowing, such as the recovery rate[38] (Box 1). However, various studies indicate that the performance of EWS metrics is reduced with more complex noise regimes[39,40] (but see refs. 41,42). Another limitation in EWSs lies in the assumptions necessary for interpreting EWSs, such as the premise that a system should be in a stable equilibrium before a tipping point occurs. In real systems, the distinction between a stable state and a long transient state may not be clear[43–45]. In addition, the rate at which the system approaches the tipping point must be sufficiently slow relative to the internal dynamics of the system (compare ref. 46), but these relative rates are typically unknown.

If EWSs are robustly detected from data, confirming slowing and resilience loss, this cannot yet be interpreted as approaching a tipping point. Slowing occurs in dynamical systems not only before tipping points (for example, fold bifurcations; Box 1) but also before other bifurcation points that are qualitatively different, and before smooth and reversible changes with no bifurcations (Fig. 1). Such other bifurcations include gradual and reversible transitions from stable equilibria to unstable equilibria (so-called transcritical bifurcations or non-catastrophic transitions) and transitions from stable equilibria to periodic oscillations and vice versa (so-called Hopf bifurcations)[12,39]. Hence, slowing and EWSs also occur in models that do not include tipping points and positive feedbacks when changing their control parameters (Fig. 1; compare refs. 47,48). In systems without tipping points, spurious EWSs may arise from stochastic fluctuations or disturbances[49,50], climate oscillations[51], self-organization and emerging chaos[52,53], (consumer-resource) cascades, such as the increase in

https://doi.org/10.1038/s41558-025-02328-8

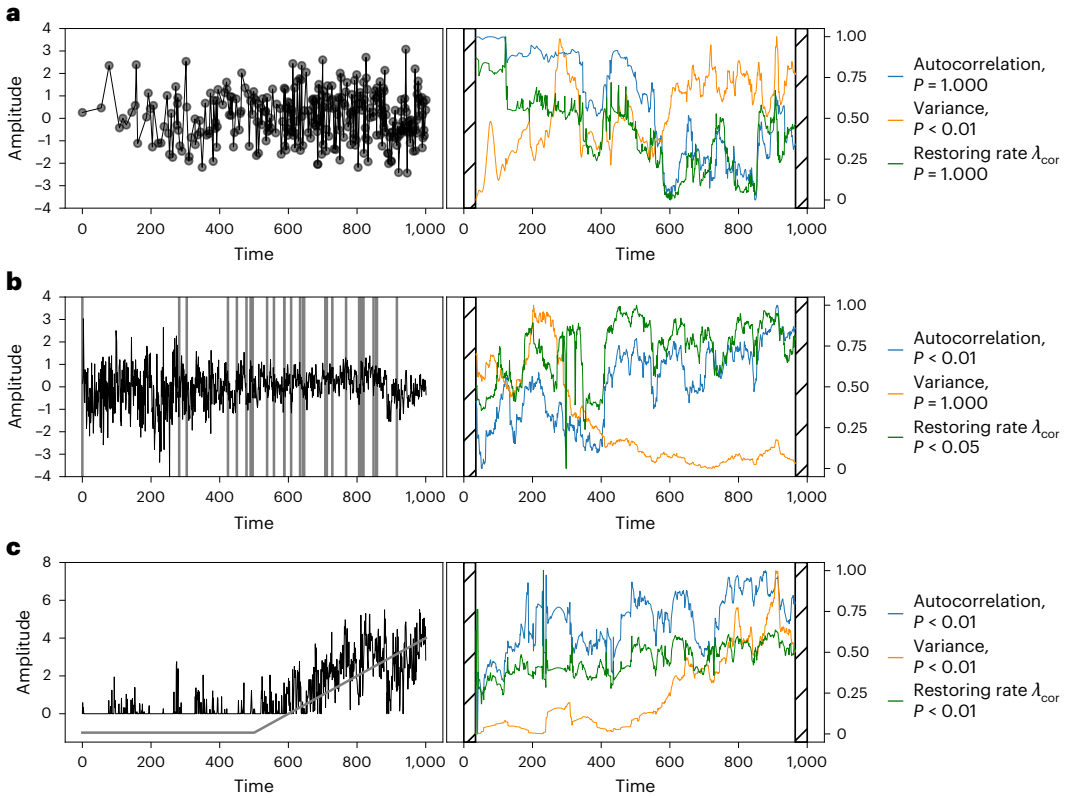


Fig. 2 | Examples for spuriously significant positive trends in EWS metrics produced from noise. Spurious increase in EWS metrics can arise due to increasing number of data points, increasing number of time series or non-stationary response and could thereby be wrongly attributed to critical slowing or decreased resilience if not accounted for. **a**, Interpolation over increasing number of data points with time can lead to increase in variance. **b**, Averaging over increasing number of time series can lead to increase in autocorrelation and corrected restoring rate λ_{cor}. Grey vertical lines mark the start point of each time series (all end at time point 1,000). **c**, A time series with a piecewise linear trend with a constant value of −1 unit for the first half of the time series and an increase from −1 to 4 units for the second half (indicated by the grey line), which is recorded only above a certain threshold (here '0'), can show increase in autocorrelation, variance and corrected restoring rate λ_{cor} (ref. 8). See Methods and Supplementary Fig. 1 for details.

rainfall due to forest transpiration upwind[54], and spatial pattern formation with associated multi-stability[18,28,55,56]. Thus, since slowing down can also occur in models without tipping points, this phenomenon alone cannot be used to robustly identify whether a tipping point is approached[12,39]. Note that even if a tipping point is approached, classical EWSs cannot be used to predict when exactly the tipping point will be crossed because they do not provide absolute distance values towards it (for example, ref. 13).

Statistical challenges for detecting EWSs

EWSs have been detected from observational data, including palaeoclimate reconstructions (for example, refs. 7,57–59), instrumental datasets (for example, ref. 8) and remote-sensing data (for example, refs. 9,60,61). In this section, we identify and reflect on pitfalls associated with the detection of EWSs in real data (Box 1).

The standard procedure is to assess the statistical significance of positive trends in EWS measures (Box 1) to determine whether the observed trend reflects expected random variations on the basis of the underlying statistical properties of the time series (compare ref. 62). For this, a null model must be defined. To ensure a meaningful result, it is important that this null model reflect relevant data properties of the system in the absence of critical slowing (for example, ref. 63).

For example, increasing data availability over time can increase the variance when more data points are available[64] (Fig. 2a) or decrease the variance when averaging over an increasing number of time series for spatial composites reduces the noise level (Fig. 2b; similar to ref. 61). In both cases, this results in an opposite trend in the lag-1 autocorrelation and corrected restoring rate λ_{cor} (with decreasing lag-1 autocorrelation and corrected restoring rate λ_{cor} when variance increases, and vice versa; Fig. 2). This underlines the sensitivity of EWS metrics to changes in data availability and quality, which needs to be accounted for when testing statistical significance.

Data products often rely on observations that vary in quantity and observation systems over time and involve processing to homogenize noisy and irregular data. Examples include variations in the density and type of palaeoclimate data[65] and transitions from palaeoclimate proxies to instrumental data or from bucket-based to more accurate buoy-based water temperature measurements (for example, ref. 66). In all cases, such changes can alter statistical properties of resulting data products (as in the example in Fig. 2a). Observational data products, such as EN4.2.1, HadSST4 or VODCA (for salinity, temperature and vegetation variables, respectively), are produced using observations with increasing availability over time, with varying types of measurement techniques, usually becoming increasingly more accurate with time[61]. Processing techniques employed to homogenize various sources of data affect higher-order statistics of resulting data products[64]. For example, in the EN4.2.1 dataset, grid cells are filled with climate background information (mean of period 1971–2000) whenever there are no data available (usually for earlier periods). This can result in increasing variance and autocorrelation in more recent periods when more data are available and, if unaccounted for, spurious positive trends of EWS metrics[64]. Likewise, data assimilation techniques, which utilize climate model data as prior information that is updated by proxy data, will lead to increase in variance when and where proxy data are available[67]. When testing for significant positive trends in EWS metrics, it is thus important to consider that these effects can occur in various combinations,

https://doi.org/10.1038/s41558-025-02328-8

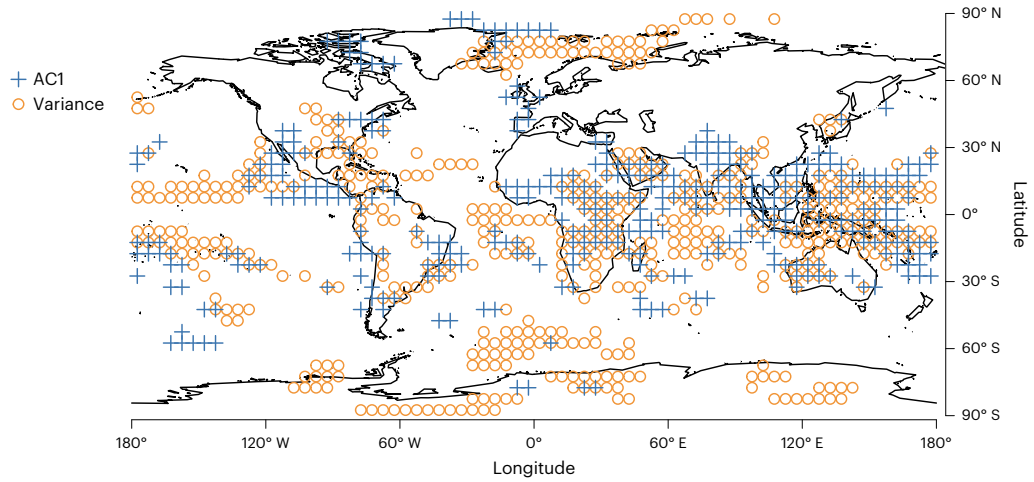


Fig. 3 | Spuriously significant positive trends in EWS metrics in a global climate model simulation. Trends of EWS metrics estimates in the MPI-ESM-P model simulation for the past millennium (past1000) are calculated for each grid point of the 2-m air-temperature field. Shown are the locations with locally statistically significant ($P < 0.05$) increase in EWS metrics, that is, in variance (circles) and autocorrelation (AC1, crosses). The statistics were calculated over a 150-year period covering 1120–1270 CE and thus encompassing one of the strongest volcanic eruptions of the Holocene epoch, the 1257 CE Samalas eruption in modern-day Indonesia. Local significant EWS trends in AC1 and variance cover 20% and 24% of Earth's surface area, respectively. However, these results are still overall not statistically significant when accounting for spatial autocorrelation ($P > 0.05$ field significance; Methods and Supplementary Fig. 2).

and inferences about critical slowing or resilience loss cannot be made without accounting for these effects.

Furthermore, changes in autocorrelation and variability unrelated to critical slowing can arise from mean climate changes (for example, refs. 68–71) or perturbations such as a random volcanic event (for example, a related temperature decrease) and can result in spuriously detected positive trends of EWSs metrics (Fig. 3 and Supplementary Fig. 2). This can be addressed by removing or otherwise considering the forced response[41] or, in case of intermittent perturbations, by having a more general stochastic surrogate model with intermittency[72].

In addition to accounting for temporal autocorrelation in significance testing, one also has to consider the prevalent spatial autocorrelation in climate fields when attempting to detect critical slowing in individual grid cells of a spatial dataset. For example, testing each grid cell of a global climate field separately for positive trends in EWS metrics with a confidence level of $P = 0.05$ will often result in more than 5% false positives, depending on the degree of spatial autocorrelation[73]. As an example, in the surface temperature field of a CMIP5 (Coupled Model Intercomparison Project Phase 5) climate model simulation, a 13% threshold is required in periods with weak volcanic activity, whereas for the period with a strong volcanic eruption, the threshold increases to 30% (Supplementary Fig. 2 and Methods) due to increasing spatial autocorrelation. Consequently, continuous areas of grid cells showing seemingly significant positive trends in EWS metrics could be erroneously interpreted as indicating critical slowing.

In palaeoclimate records, secondary processes change the statistical properties of measured proxies. For example, biological mixing in sediments or isotope diffusion in ice cores reduces variance and increases autocorrelation[74]. Changes in parameters of the recording process, such as the rate of sediment or snow accumulation, are often linked to changes in the mean climate and modulate the time-series properties by affecting the time window over which the mixing occurs, as well as influencing the noise level[75]. Finally, in nonlinear observational systems such as melt records in ice cores[7], where crossing the melting point is required to record any signal, mean climate changes can result in increases in autocorrelation and variance (Fig. 2c).

The estimation of EWS metrics usually involves a range of methodological and parameter choices such as for detrending and removal of the seasonal cycle. The number of choices has to be limited and based on independent data or information sources to avoid an inflation of false positives due to the multiple testing effect[76]. For that, prior knowledge on the timescale of perturbations[41] and the intrinsic timescale of the system, that is, on which it responds to the perturbations, is important, for example, when choosing the window size for analysis[63]. In many cases, this knowledge is not available, and therefore, detecting critical slowing through trends in EWS measures remains a challenging statistical problem. Successful detection depends strongly on the length of the time series, the amount of noise and the rate at which the tipping point is approached (for example, refs. 77,78).

Bayesian framework for more robust early warning

Studies generally interpret EWSs detected from observational data as indication that a tipping point is approaching. The probability of encountering EWSs when approaching a tipping point is relatively high because of associated critical slowing (Box 1). However, this is easily confused with the probability of a tipping point approaching under the condition that EWSs are observed. This latter probability could be much lower, for example, due to the existence of multiple alternative explanations for the presence of positive EWSs, besides chance (Fig. 4).

To further illustrate this, a comparable error would be to conflate the probability of a car alarm going off given the condition that a burglary occurs, with the probability that a burglary is happening given the condition that the car alarm goes off. The first probability will be relatively high, as the breaking of the window glass will be sensed by the alarm. However, the latter probability will be relatively low, as events more likely than a burglary, such as a passing truck or a thunderclap, could have the same effect. Conflating these probabilities is called 'transposing the conditional', which results in biased assessment of statistical significance[79–82]. This issue has also been observed in studies employing EWS theory to detect approaching tipping points[83].

To overcome these challenges, we advocate the use of a Bayesian framework[82] to consider the prior probabilities of the hypothesis, in our case the prior evidence of approaching tipping points besides EWSs as well as the alternative explanations for positive trends of EWS measures in the data (Fig. 4). Information on the occurrence and frequency of tipping events independent of the EWSs is required to estimate the prior probability of a tipping point approaching, $P(\mathrm{TP})$ (Fig. 4). Important information sources are model simulations, either

https://doi.org/10.1038/s41558-025-02328-8

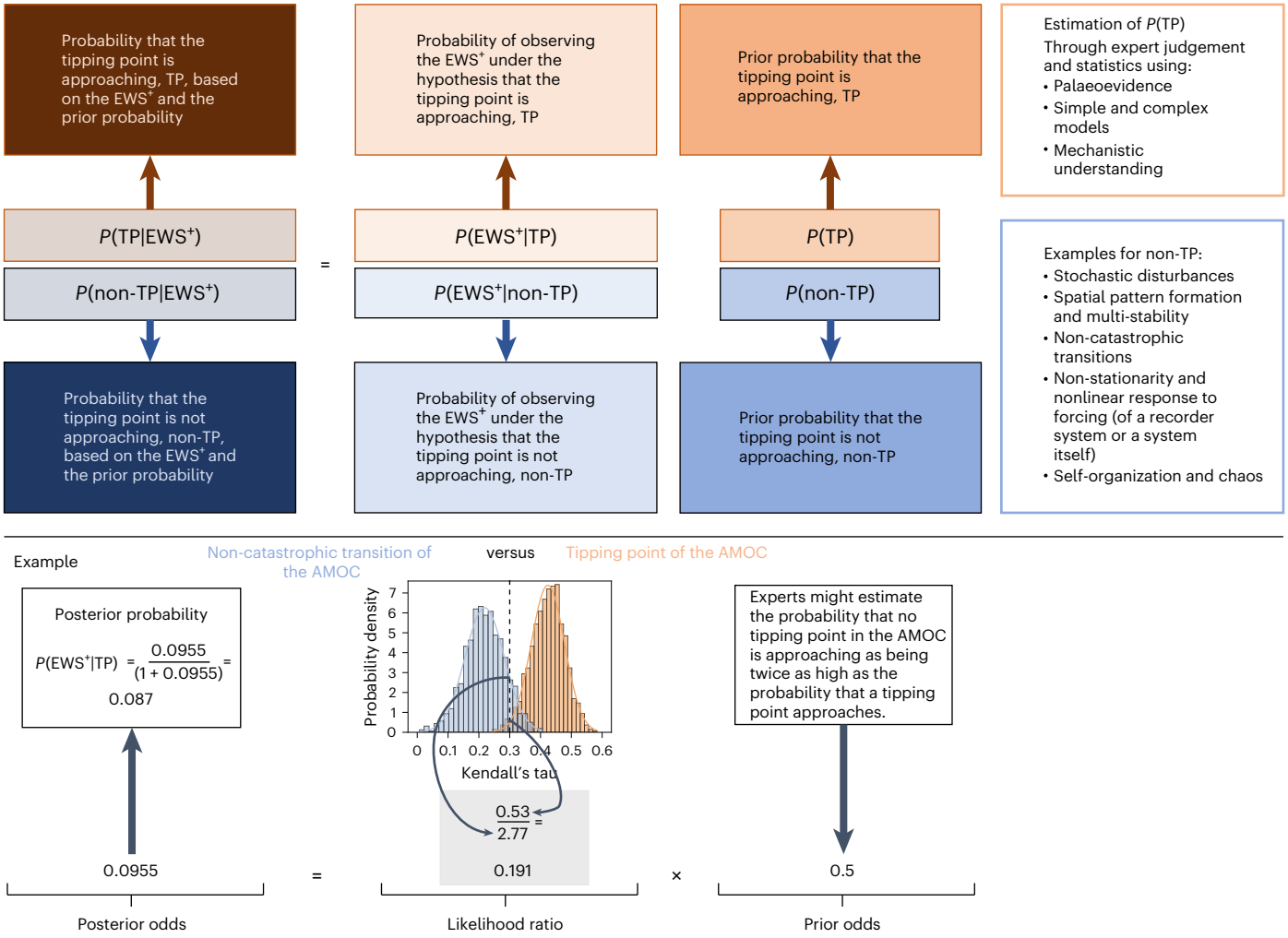


Fig. 4 | Bayesian framework for estimating the probability of approaching a tipping point based on observational data via EWSs. TP is the hypothesis that a tipping point is approaching, and non-TP is the hypothesis that no tipping point is approaching and instead includes potential alternative explanations for EWSs. *P*(TP) is the probability for the TP hypothesis and *P*(non-TP) for the non-TP hypothesis (or hypotheses, since non-TP should be the union of all the possible alternative hypotheses; Methods). Listed in the box on the right are examples for non-TP given in this Perspective. *P*(EWS$^+$|TP) is the probability that the EWS is found under the hypothesis that the tipping point is approaching. *P*(EWS$^+$|non-TP) is the probability that the EWS is found under the assumption that no tipping point is approaching. What is usually tested for is only the probability of observing EWSs at least as extreme as the one observed solely by chance, that is, considering only random fluctuations of data properties as a non-TP hypothesis. The left-hand side of the equation represents the posterior odds for tipping to occur. The probability of a forthcoming tipping event, *P*(TP|EWS$^+$), can be directly estimated from the posterior odds ratio, *P*(TP|EWS$^+$)/*P*(non-TP|EWS$^+$), but will be overestimated if *P*(non-TP|EWS$^+$) is not exhaustive. The bottom of the figure describes an example application of the framework using the AMOC box model utilized for Fig. 1 and based on ref. 17. See text for further explanation and Methods for details.

systematically analysing and interpreting recent ESM ensembles (compare ref. 19) or specific model studies investigating tipping mechanisms in detail (compare ref. 17). A complementary source to estimate the prior probability of tipping points is palaeodata[84]. Regional analogue situations, such as the Greenland climate during past interglacials[85], or spatial analogues (similar climate changes but at different locations) can provide information about the probability of tipping points. Finally, estimation of the prior probability of an approaching tipping point might also employ expert judgement, which integrates the evidence from models, palaeodata and mechanistic understanding (Fig. 4). Although these probabilities are also subject to high uncertainty, including and explicitly stating the prior probability allows for transparency and adds information as there are some tipping points with lower or higher prior probabilities[3,5]. If the prior probability is lower, stronger evidence derived from EWS measures through the likelihood ratio, *P*(EWS$^+$|TP)/*P*(EWS$^+$|non-TP), is necessary to yield a high probability of an approaching tipping point based on the observed data, *P*(TP|EWS$^+$) (Fig. 4).

Studies searching for EWSs using observational data often consider only expected random variations in the underlying statistical properties of the time series as alternative explanations for observing positive trends in EWS metrics besides the TP hypothesis. However, more parsimonious alternative explanations exist, potentially resulting in positive trends in EWS metrics, some of which are discussed in this Perspective (Fig. 4). For example, slowing can precede a wide range of impending transitions, hence can also occur for transitions without tipping points. This implies that one also has to estimate the probability of positive trends in EWS metrics under the hypothesis that no tipping point is approaching, *P*(EWS$^+$|non-TP), and consider the prior odds, *P*(TP)/*P*(non-TP). While it might be impossible to account for all possible non-TP hypotheses, already including the most likely alternative explanations, besides random variations, will lead to more accurate estimates of *P*(TP|EWS$^+$).

To illustrate the framework, let us assume we have observed a specific trend in EWS metric from observations, here chosen as an arbitrary Kendall's tau of 0.3 (Fig. 4). Following the frequentist approach,

this observed value would be compared with realizations of a stationary stochastic process and considered statistically significant ($P < 0.05$ in our example; Methods). This result might be misinterpreted as indicating a tipping probability greater than 95% (ref. 86). In the ideal case, we have a realistic model for the system approaching a tipping point alongside one or multiple alternative models. Such a model-based approach for tipping-point detection using EWSs was proposed by ref. 83.

For our toy example, we use a simplified AMOC model[17] where varying parameter values allow simulation of both TP and non-TP scenarios. This allows us to calculate the probabilities for observing the trend in the EWS metric (Kendall's tau of 0.3) under both hypotheses and the likelihood ratio. In our example, we simplify by using realizations from two fixed parameters (thus assuming a fixed expected time to tipping and to a non-catastrophic transition), resulting in a likelihood ratio of 0.191. This would mean that the trend in the EWS metric is about five times more likely under the model where no tipping point is approaching compared with the model where a tipping point is approaching. If the scenario without a tipping point is twice as likely as the TP scenario, that is, prior odds of 0.5, then the resulting posterior odds for an approaching tipping point would be 0.0955, corresponding to a tipping probability of 0.087 (equation (6)). Note that this is a hypothetical calculation with an arbitrary choice of the observed value and parameters for illustration, and no inferences can be made from it. As prior probabilities are often unknown and have to be estimated, the framework adds more complexity to applications. However, it makes these unknowns transparent and reveals what needs to be better understood to accurately assess the risk.

Recommendations for future research

Crossing climate tipping points can lead to catastrophic impacts. To prevent harm to humanity, one might prefer not to accept any risk and to sound a false alarm rather than to fail in predicting an imminent catastrophe, according to the precautionary principle. However, here we argue that the probability of a false alarm could be reduced through more accurate assessments of tipping probabilities. When detecting approaching tipping points, the amount of observational data will always be limited (as in reality we can typically observe only one trajectory of an Earth system element); therefore, incorporating physical reasoning and prior knowledge into the statistical analysis is necessary[82].

In general, mechanistic understanding of the processes associated with the tipping point under study is required. For example, the relevant intrinsic timescales of the tipping element need to be known to inform parameter choices for statistical analysis, such as the window size. In addition, the connections between relevant timescales and mechanisms help to explain observations made. For example, resilience loss of Amazon rainforest seems to be associated with drought and human disturbance, pointing towards biological mechanisms related to water stress and land use[9].

Furthermore, the power of the data for discriminating between the TP and non-TP hypotheses will be maximized by an optimal combination of well-tested EWS metrics, as well as including processes in the probability model that might bias the EWS metrics, such as changes in the amount of data. An accurate probability model also accounts for temporal and spatial correlations[87] as well as multiple testing. This is necessary regardless of whether a frequentist or Bayesian approach is applied.

To further improve the detection of approaching tipping points, EWSs based on a mechanistic understanding of the physical or biological mechanisms, going beyond the statistical EWS metrics, should be developed[25,88]. These could include multiple variables, such as ice-sheet height and meltwater flux, and involve specifically designed observation systems of a tipping-prone system, such as the RAPID array observing the overturning circulation[89].

Recently, AMOC-induced freshwater transport at the southern boundary of the Atlantic has been proposed as a physics-based EWS on the basis of an AMOC tipping event forced in the community Earth system model[25]. Future studies also need to consider the probability that the same signal would occur in the absence of tipping under similar forcing. According to the Bayesian framework, this influences the evidence for approaching AMOC tipping obtained from the EWS as higher probabilities for the alternative hypothesis lower the likelihood ratio for a tipping to occur (Fig. 4).

Current community modelling efforts, such as the Tipping Point Model Intercomparison Project, will provide a set of climate model simulations exploring the tipping mechanisms and critical thresholds of Earth system components. Such model simulations will, on one hand, provide probabilities for future crossing of tipping points independent of EWSs and can thus act as prior information for probability assessments. On the other hand, they will also allow us to improve the use of EWSs via their simulated linkage between EWSs and approaching tipping events. This can be done directly such as recently done[25] or by using the models for estimating the likelihood ratio of the EWSs (Fig. 4). While it will be challenging to create such large ensembles of simulations usually needed in this procedure, methods targeted to increase the ensemble size such as emulators or machine learning, and other methods to efficiently approximate likelihoods from complex simulators[90], will prove useful. When formulating the hypothesis to test for, one needs to consider how much time is left to the tipping event. Thus, when producing a model ensemble, one might choose either a prior distribution of the parameter that determines the time to tipping or an ensemble of simulations where the tipping occurs at different times. In cases where changing mean state can be neglected, selecting different periods relative to the tipping event allows for the inclusion of a prior on the time to tipping without requiring larger ensembles.

Here we considered tipping points to be associated exclusively with fold bifurcations. However, one might want to use EWSs to additionally detect other types of bifurcations. This can similarly be achieved by considering the individual probabilities for each type of bifurcation.

After obtaining the probability of a forthcoming tipping event on the basis of the observed data, one needs to decide what probability is deemed 'acceptable' to raise alarm. Making this decision involves considering the size of the risk, which is the product of the probability and the consequence. In climate tipping studies, raising alarm could be already warranted in case of finding a relatively low probability of an imminent tipping event to occur. To address this challenge and include the consequences in decision making, physical climate story lines[91], climate model simulations in which certain phenomena such as tipping points are artificially prescribed, are a promising approach[92].

Finally, to reduce the uncertainties in risk assessment of tipping points, all lines of evidence from observational datasets, palaeoclimate information, different climate models and expert judgement should be combined. Here Bayesian networks[93] can provide a tool to integrate these different information sources into a transparent and reproducible risk assessment.

Imminent tipping is often put forward as an important motivation for urgent climate action (for example, refs. 4,94). Such strong claims of approaching tipping clearly demand robust evidence. If tipping points that were predicted to be imminent do not occur as expected, this could undermine communication efforts on the necessity of mitigation. We argue that the robustness of this evidence can be increased by carefully considering the theoretical and statistical challenges associated with the use of EWSs. In light of the potentially dangerous outcomes of crossing tipping points, the precautionary principle should apply until improved probabilistic assessments are available that result in robust risk assessments for informed decision making. However, regardless of whether climate tipping points are imminent, immediate and transformative climate action remains essential[95]. The ongoing

and escalating impacts of gradual climate change on nature and society, exemplified by the increasing frequency and severity of extreme weather events[5], including devastating floods, heatwaves, droughts and wildfires, demand urgent and decisive measures, independent of uncertainties surrounding tipping points.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41558-025-02328-8.

References

1. Scheffer, M., Carpenter, S., Foley, J. A., Folke, C. & Walker, B. Catastrophic shifts in ecosystems. *Nature* **413**, 591–596 (2001).
2. Lenton, T. M. et al. Tipping elements in the Earth's climate system. *Proc. Natl Acad. Sci. USA* **105**, 1786–1793 (2008).
3. Armstrong McKay, D. I. et al. Exceeding 1.5 °C global warming could trigger multiple climate tipping points. *Science* **377**, eabn7950 (2022).
4. Lenton, T. M. et al. *The Global Tipping Points Report 2023* (Univ. Exeter, 2023).
5. IPCC *Climate Change 2021: The Physical Science Basis* (eds Masson-Delmotte, V. et al.) (Cambridge Univ. Press, 2021).
6. Scheffer, M. et al. Early warning signals for critical transitions. *Nature* **461**, 53–59 (2009).
7. Boers, N. & Rypdal, M. Critical slowing down suggests that the western Greenland Ice Sheet is close to a tipping point. *Proc. Natl Acad. Sci. USA* **118**, e2024192118 (2021).
8. Boers, N. Observation-based early-warning signals for a collapse of the Atlantic meridional overturning circulation. *Nat. Clim. Change* **11**, 680–688 (2021).
9. Boulton, C. A., Lenton, T. M. & Boers, N. Pronounced loss of Amazon rainforest resilience since the early 2000s. *Nat. Clim. Change* **12**, 271–278 (2022).
10. Ditlevsen, P. & Ditlevsen, S. Warning of a forthcoming collapse of the Atlantic meridional overturing circulation. *Nat. Commun.* **14**, 4254 (2023).
11. Boettiger, C. & Hastings, A. Early warning signals and the prosecutor's fallacy. *Proc. R. Soc. B* **279**, 4734–4739 (2012).
12. Kéfi, S., Dakos, V., Scheffer, M., Van Nes, E. H. & Rietkerk, M. Early warning signals also precede non-catastrophic transitions. *Oikos* **122**, 641–648 (2013).
13. Ben-Yami, M., Morr, A., Bathiany, S. & Boers, N. Uncertainties too large to predict tipping times of major Earth system components from historical data. *Sci. Adv.* **10**, eadl4841 (2024).
14. Van Nes, E. H. et al. What do you mean, 'tipping point'? *Trends Ecol. Evol.* **31**, 622–632 (2016).
15. Ashwin, P., Wieczorek, S., Vitolo, R. & Cox, P. Tipping points in open systems: bifurcation, noise-induced and rate-dependent examples in the climate system. *Phil. Trans. R. Soc. A* **370**, 1166–1184 (2012).
16. Titz, S., Kuhlbrodt, T., Rahmstorf, S. & Freudel, U. On freshwater-dependent bifurcations in box models of the interhemispheric thermohaline circulations. *Tellus A* **54**, 89–98 (2002).
17. Lux, K., Ashwin, P., Wood, R. & Kuehn, C. Assessing the impact of parametric uncertainty on tipping points of the Atlantic meridional overturing circulation. *Environ. Res. Lett.* **17**, 075002 (2022).
18. Lohmann, J., Dijkstra, H. A., Jochum, M., Lucarini, V. & Ditlevsen, P. D. Multistability and intermediate tipping of the Atlantic ocean circulation. *Sci. Adv.* **10**, eadi4253 (2024).
19. Drijfhout, S. et al. Catalogue of abrupt shifts in Intergovernmental Panel on Climate Change climate models. *Proc. Natl Acad. Sci. USA* **112**, E5777–E5786 (2015).
20. Bathiany, S., Notz, D., Mauritsen, T., Raedel, G. & Brovkin, V. On the potential for abrupt Arctic winter sea ice loss. *J. Clim.* **29**, 2703–2719 (2016).
21. Stommel, H. Thermohaline convection with two stable regimes of flow. *Tellus* **13**, 224–230 (1961).
22. Cessi, P. A. A simple box model of stochastically forced thermohaline flow. *J. Phys. Oceanogr.* **24**, 1911–1920 (1994).
23. Liu, W., Liu, Z. & Brady, E. C. Why is the AMOC monostable in coupled general circulation models? *J. Clim.* **27**, 2427–2443 (2014).
24. Jackson, L. C. et al. Understanding AMOC stability: the North Atlantic Hosing Model Intercomparison Project. *Geosci. Model Dev.* **16**, 1975–1995 (2023).
25. Van Westen, R. M., Kliphuis, M. & Dijkstra, H. A. Physics-based early warning signal shows that AMOC is on tipping course. *Sci. Adv.* **10**, eadk1189 (2024).
26. Gildor, H. & Tziperman, E. Physical mechanisms behind biogeochemical glacial–interglacial CO_2 variations. *Geophys. Res. Lett.* **28**, 2421–2424 (2001).
27. Alkhayoun, H. M., Ashwin, P., Jackson, L. C., Quinn, C. & Wood, R. A. Basin bifurcation, oscillatory instability and scale induced thresholds for the Atlantic meridional overturning circulation in a global box model. *Proc. R. Soc. A* **475**, 20190051 (2019).
28. Rietkerk, M. et al. Evasion of tipping in complex systems through spatial pattern formation. *Science* **374**, eabj0359 (2021).
29. Bastiaansen, R., Dijkstra, H. A. & Von der Heydt, A. S. Fragmented tipping in a spatially heterogeneous world. *Environ. Res. Let.* **17**, 045006 (2022).
30. Laepple, T. et al. Regional but not global temperature variability underestimated by climate models at supradecadal timescales. *Nat. Geosci.* **16**, 958–966 (2024).
31. Dekker, M. M., von der Heydt, A. S. & Dijkstra, H. A. Cascading transitions in the climate system. *Earth Syst. Dyn.* **9**, 1243–1260 (2018).
32. Klose, A. K., Karle, V., Winkelmann, R. & Donges, J. F. Emergence of cascading dynamics in interacting tipping elements of ecology and climate. *R. Soc. Open Sci.* **7**, 200599 (2020).
33. Klose, A. K., Wunderling, N., Winkelmann, R. & Donges, J. F. What do we mean, 'tipping cascade'? *Environ. Res. Let.* **16**, 125011 (2021).
34. Liu, T. et al. Teleconnections among tipping elements in the Earth system. *Nat. Clim. Change* **13**, 67–74 (2023).
35. Wunderling, N. et al. Global warming overshoots increase risks of climate tipping cascades in a network model. *Nat. Clim. Change* **13**, 75–82 (2023).
36. Wunderling, N., Donges, J. F., Kurths, J. & Winkelmann, R. Interacting tipping elements increase risk of climate domino effects under global warming. *Earth Syst. Dyn.* **12**, 601–619 (2021).
37. Wunderling, N. et al. Climate tipping point interactions and cascades: a review. *Earth Syst. Dyn.* **15**, 41–74 (2024).
38. Held, H. & Kleinen, T. Detection of climate system bifurcations by degenerate fingerprinting. *Geophys. Res. Let.* **31**, L23207 (2004).
39. Weinans, E., Ouax, R., Van Nes, E. H. & Van de Leemput, I. A. Evaluating the performance of multivariate indicators of resilience loss. *Sci. Rep.* **11**, 9148 (2021).
40. Laitinen, V. & Lahti, L. in *Computational Methods in Systems Biology* (eds Petre, I. & Păun, A.) 259–274 (Springer, 2022).
41. Benson, V. et al. Measuring tropical rainforest resilience under non-Gaussian disturbances. *Environ. Res. Lett.* **19**, 024029 (2024).
42. Morr, A. & Boers, N. Detection of approaching critical transitions in natural systems driven by red noise. *Phys. Rev. X* **14**, 021037 (2024).
43. Dakos, V. et al. Tipping point detection and early warnings in climate, ecological, and human systems. *Earth Syst. Dyn.* **15**, 1117–1135 (2024).

44. Hastings, A. et al. Transient phenomena in ecology. *Science* **361**, eaat6412 (2018).
45. Sánchez-Pinillos, M., Dakos, V. & Kéfi, S. Ecological dynamic regimes: a key concept for assessing ecological resilience. *Biol. Conserv.* **289**, 110409 (2024).
46. Dablander, F. et al. Overlapping timescales obscure early warning signals of the second COVID-19 wave. *Proc. R. Soc. B* **289**, 20211809 (2022).
47. Bury, T. M., Bauch, C. T. & Anand, M. Detecting and distinguishing tipping points using spectral early warning signals. *J. R. Soc. Interface* **17**, 20200482 (2020).
48. Bury, T. M. et al. Deep learning for early warning signals of tipping points. *Proc. Natl Acad. Sci. USA* **118**, e2106140118 (2021).
49. Ghil, M. in *Encyclopedia of Global Environmental Change* Vol. 1 (eds Munn, T. E. et al.) 544–549 (John Wiley & Sons, 2002).
50. Ditlevsen, P. D. & Johnsen, S. J. Tipping points: early warning and wishful thinking. *Geophys. Res. Let.* **37**, L19703 (2010).
51. Dijkstra, H. A. & Ghil, M. Low-frequency variability of the large-scale ocean circulation: a dynamical systems approach. *Rev. Geophys.* **43**, RG3002 (2005).
52. Bascompte, J. & Solé, R. V. Rethinking complexity: modelling spatiotemporal dynamics in ecology. *Trends Ecol. Evol.* **10**, 361–366 (1995).
53. Solé, R. & Bascompte, J. *Self-Organization in Complex Ecosystems* (Princeton Univ. Press, 2006).
54. Staal, A. et al. Forest-rainfall cascades buffer against drought across the Amazone. *Nat. Clim. Change* **8**, 539–543 (2018).
55. Meron, E. *Nonlinear Physics of Ecosystems* (CRC Press, 2015).
56. Bastiaansen, R. et al. Multistability of model and real dryland ecosystems through spatial self-organization. *Proc. Natl Acad. Sci. USA* **115**, 11256–11261 (2018).
57. Dakos, V. et al. Slowing down as an early warning signal for abrupt climate change. *Proc. Natl Acad. Sci. USA* **105**, 14308–14312 (2008).
58. Boers, N. Early-warning signals for Dansgaard–Oeschger events in a high-resolution ice core record. *Nat. Commun.* **9**, 2556 (2018).
59. Michel, S. L. et al. Early warning signal for a tipping point suggested by a millennial Atlantic multidecadal variability reconstruction. *Nat. Commun.* **13**, 5176 (2022).
60. Verbesselt, J. et al. Remotely sensed resilience of tropical forests. *Nat. Clim. Change* **6**, 1028–1031 (2016).
61. Smith, T. et al. Reliability of resilience estimation based on multi-instrument time series. *Earth Syst. Dyn.* **14**, 173–183 (2023).
62. Rocha, J. C. Ecosystems are showing symptoms of resilience loss. *Environ. Res. Let.* **17**, 065013 (2022).
63. Dakos, V. et al. Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data. *PLoS ONE* **7**, e41010 (2012).
64. Ben-Yami, M., Skiba, V., Bathiany, S. & Boers, N. Uncertainties in critical slowing down indicators of observation-based fingerprints of the Atlantic overturing circulation. *Nat. Commun.* **14**, 8344 (2023).
65. Anchukaitis, K. J. & Smerdon, J. E. Progress and uncertainties in global and hemispheric temperature reconstructions of the Common Era. *Quat. Sci. Rev.* **286**, 107537 (2022).
66. Kennedy, J. J., Rayner, N. A., Atkinson, C. P. & Killick, R. E. An ensemble data set of sea surface temperature change from 1850: the Met Office Hadley Centre HadSST. 4.0. 0.0 data set. *J. Geophys. Res. Atmos.* **124**, 7719–7763 (2019).
67. King, J. M. et al. A data assimilation approach to last millennium temperature field reconstruction using a limited high-sensitivity proxy network. *J. Clim.* **34**, 7091–7111 (2021).
68. Huntingford, C., Jones, P. D., Livina, V. N., Lenton, T. M. & Cox, P. M. No increase in global temperature variability despite changing regional patterns. *Nature* **500**, 327–330 (2013).
69. Bathiany, S. et al. Statistical indicators of Arctic sea ice stability—prospects and limitations. *Cryosphere* **10**, 1631–1645 (2016).
70. Lenton, T. M. et al. Observed trends in the magnitude and persistence of monthly temperature variability. *Sci. Rep.* **7**, 5940 (2017).
71. Rehfeld, K., Münch, T., Ho, S. L. & Laepple, T. Global patterns of declining temperature variability from the Last Glacial Maximum to the Holocene. *Nature* **554**, 356–359 (2018).
72. Lovejoy, S. & Varotsos, C. Scaling regimes and linear/nonlinear responses of last millennium climate to volcanic and solar forces. *Earth Syst. Dyn.* **7**, 133–150 (2016).
73. Livezey, R. E. & Chen, W. Y. Statistical field significance and its determination by Monte Carlo techniques. *Mon. Weather Rev.* **111**, 46–59 (1983).
74. Dolman, A. M. & Laepple, T. Sedproxy: a forward model for sediment-archived climate proxies. *Clim. Past* **14**, 1851–1868 (2018).
75. Hirsch, N. et al. Stratigraphic noise and its potential drivers across the plateau of Dronning Maud Land, East Antarctica. *Cryosphere* **17**, 4207–4221 (2023).
76. Davis, R. E. Predictability of sea surface temperature and sea level pressure anomalies over the North Pacific Ocean. *J. Phys. Oceanogr.* **6**, 249–266 (1976).
77. Clements, C. F., Drake, J. M., Griffiths, J. I. & Ozgul, A. Factors influencing the detectability of early warning signals of population collapse. *Am. Nat.* **186**, 50–58 (2015).
78. Dablander, F., Pichler, A., Cika, A. & Bacilieri, A. Anticipating critical transitions in psychological systems using early warning signals: theoretical and practical considerations. *Psychol. Methods* **28**, 765–790 (2023).
79. Thompson, W. C. & Schumann, E. L. Interpretation of statistical evidence in criminal trials: the prosecutor's fallacy and the defense attorney's fallacy. *Law Hum. Behav.* **11**, 167–187 (1987).
80. Evett, I. W. Avoiding the transposed conditional. *Sci. Justice* **2**, 127–131 (1995).
81. Neath, A. A. Statistical inference, statistics education, and the fallacy of the transposed conditional. In *Proc. Joint Statistical Meeting* 3348–3350 (2010).
82. Shepherd, T. G. Bringing physical reasoning into statistical practice in climate-change science. *Climatic Change* **169**, 2 (2021).
83. Boettiger, C. & Hastings, A. Quantifying limits to detection of early warning for critical transitions. *J. R. Soc. Interface* **9**, 2527–2539 (2012).
84. Brovkin, V. et al. Past abrupt changes, tipping points and cascading impacts in the Earth system. *Nat. Geosci.* **14**, 550–558 (2021).
85. NEEM community members. Eemian interglacial reconstructed from a Greenland folded ice core. *Nature* **493**, 489–494 (2013).
86. Greenland, S. et al. Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *Eur. J. Epidemiol.* **31**, 337–350 (2016).
87. Ives, A. R. et al. Statistical inference for trends in spatiotemporal data. *Remote Sens. Environ.* **266**, 112678 (2021).
88. Watkins, N. W., Pruessner, G., Chapman, S. C., Crosby, N. B. & Jensen, H. J. 25 years of self-organized criticality: concepts and controversies. *Space Sci. Rev.* **198**, 3–44 (2016).
89. Srokosz, M. A., Holliday, N. P. & Bryden, H. L. Atlantic overturning: new observations and challenges. *Phil. Trans. R. Soc. A* **381**, 20220196 (2023).
90. Cranmer, K., Brehmer, J. & Louppe, G. The frontier of simulation-based inference. *Proc. Natl Acad. Sci. USA* **117**, 30055–30062 (2020).
91. Shepherd, T. G. et al. Storylines: an alternative approach to representing uncertainty in physical aspects of climate change. *Climatic Change* **151**, 555–571 (2018).

92. Ritchie, P. D. L. et al. Shifts in national land use and food production in Great Britain after a climate tipping point. *Nat. Food* **1**, 76–83 (2020).
93. Sahlin, U., Helle, I. & Perepolking, D. 'This is what we don't know': treating epistemic uncertainty in Bayesian networks for risk assessment. *Integr. Environ. Assess. Manage.* **17**, 221–232 (2021).
94. Lenton, T. M. et al. Climate tipping points—too risky to bet against. *Nature* **575**, 592–595 (2019).
95. Morrison, T. H. et al. Radical interventions for climate-impacted systems. *Nat. Clim. Change* **12**, 1100–1106 (2022).
96. Patterson, A. C., Strang, A. G. & Abbott, K. C. When and where can we expect to see early warning signals in multispecies systems approaching tipping points: insights from theory. *Am. Nat.* **198**, E12 (2021).
97. Morr, A., Boers, N. & Ashwin, P. Internal noise interference to warnings of tipping points in generic multi-dimensional dynamical systems. *SIAM J. Appl. Dyn. Syst.* **23**, 2793–2806 (2024).
98. Drake, J. M. & Griffin, B. D. Early warning signals of extinction in deteriorating environments. *Nature* **467**, 456–459 (2010).
99. Ricker, W. E. Stock and recruitment. *Can. J. Fish. Aquat. Sci.* **11**, 559–623 (1954).
100. Rietkerk, M., Ketner, P., Stroosnijder, L. & Prins, H. H. Sahelian rangeland development; a catastrophe? *J. Range Manage.* **49**, 512–519 (1996).
101. Wissel, C. A. Universal law of the characteristic return time near thresholds. *Oecologia* **65**, 101–107 (1984).
102. Boettner, C. & Boers, N. Critical slowing down in dynamical systems driven by non-stationary correlated noise. *Phys. Rev. Res.* **4**, 013230 (2022).
103. Heßler, M. & Kamps, O. Bayesian on-line anticipation of critical transitions. *N. J. Phys.* **24**, 063021 (2022).
104. Clarke, J. J., Huntingford, C., Ritchie, P. D. & Cox, P. M. Seeking more robust early warning signals for climate tipping points: the ratio of spectra method (ROSA). *Environ. Res. Let.* **18**, 035006 (2023).
105. Smith, T., Traxl, D. & Boers, N. Empirical evidence for recent global shifts in vegetation resilience. *Nat. Clim. Change* **12**, 477–484 (2022).

Methods

EWS calculation and significance test

For EWS metric calculation, we run 70-year windows over the time series and linearly detrend each sub-segment. From these residuals of each window, we compute the statistical properties of interest (variance, AC1 and λ_{cor}; details in ref. 8). Kendall's tau is then employed to estimate the correlation between the resulting time series and time.

To test whether the trends in EWS metrics are statistically significant, surrogate data are generated by randomizing the phases of the Fourier transforms of the studied time series. This approach preserves the statistical characteristics of the original time series while ensuring the surrogate data are stationary. In this way, we generate an ensemble of 1,000 Fourier surrogates on which we perform the same analysis, computing the EWS metrics and the Kendall's tau as measure of the trend. The P value is derived from the ensemble as the proportion of members with a higher Kendall's tau than the original time series.

We use the code published in ref. 8 and calculate the data surrogates as in ref. 64. To calculate the λ_{cor}, ref. 8 used a linear regression for dx onto x in the equilibrium point. To reflect the eigenvalue more closely, we divide by the size of the time step and thus regress dx/dt onto x in the equilibrium point. This leads to a linear scaling of the calculated lambda and thus has no effect on our visualizations, as we scale all EWS metrics between zero and one.

The same method is used to calculate the statistically significant positive trends in EWS metrics for the MPI-ESM-P past1000 simulation covering the 850–1850 CE period. The significance is assessed for 70-year running windows on 150-year intervals with 1,000 Fourier surrogates. We show the results for the 1120–1270 CE period as the presence of the Samalas eruption in 1261 results in particularly widespread spuriously significant positive trends of EWS metrics (Fig. 3).

To calculate the spatial area (as a fraction of Earth's surface) required for field significance, we take a Monte Carlo approach based on ref. 73. We calculate the spatial area that is significantly correlated, that is, where $P < 0.05$ locally, using 10,000 fractional noise realizations with power-law type scaling behaviour, that is, $S(f) = f^{\beta}$ with $\beta = 0.5$. The distribution of locally significantly correlated spatial areas indicates the effective spatial degrees of freedom and can be used to derive the threshold needed to obtain spatial statistical significance. For a significance level $P < 0.05$, we need a spatial area greater than the 95% quantile of the distributions. We performed this analysis for two periods: 850–1000 CE, which contains only weak volcanism, and 1120–1270 CE, which includes the strong Samalas eruption towards its end (Supplementary Fig. 2a,b). The resulting distributions (Supplementary Fig. 2c) align with theoretical expectations[73] and show a higher threshold for the spatial area (30%) for the period with the Samalas eruptions than for the period with weak volcanism (13%). On the basis of the binomial distribution[73], we calculate a lower number of effective spatial degrees of freedom around 7 for the former period, compared with around 37 effective spatial degrees of freedom for the latter period. Therefore, while the volcanic forcing increases the spatial area covered by locally significant positive trends in EWS metrics, it still falls short of being considered significant if the higher spatial autocorrelation is taken into account.

AMOC model

For Fig. 1, we consider a stochastic version of the simple Stommel–Cessi AMOC model[22] in its dimensionless form as used and analysed by ref. 17.

This model is described by:

$$\mathrm{d}x = \left[\mu - x\left(1 + \eta^2\left(1 + x\right)\right)\right]\mathrm{d}t + \sigma \mathrm{d}W \tag{1}$$

where x represents a dimensionless salinity difference, and the bifurcation parameter μ represents the freshwater flux. Depending on the parameter η^2, the model can show either a smooth transition or a critical transition, with two saddle-node bifurcations (see ref. 17 for a more extensive analysis of the model). For our smooth transitions, we set $\eta^2 = 2$, and for our critical transitions, we set $\eta^2 = 7.5$. The noise is simulated with a Wiener process with a standard deviation of 0.005. We use a Euler Maruyama scheme with an integration step of 0.0025 and simulate for 200,000 time steps using the sdeint package in Python.

Generalized examples

For all generalized examples in Fig. 2, we produce a 1,000 time-units-long climate surrogate time series consisting of realizations of fractional noise with power-law type scaling behaviour with a scaling exponent β of 1. Such a scaling relationship has been shown to well approximate the behaviour of climate time series over a large range of timescales[106].

Interpolation over an increasing number of data points. For Fig. 2a, we add white noise (mean = 0, s.d. = 1) to the power-law time series (scaled to s.d. = 0.5) to mimic a typical observation that can be described as the sum of a climate signal and measurement- or sampling-induced noise. We randomly draw 300 samples from this time series (n = 1,000) with linearly increasing probability to be drawn with time (from 0 to 1 for time point 0 to 1,000). Thus, the youngest time point is always drawn. We also add the oldest time point so that the final time series covers the full period of the original time series. This results in approximately 20 times more points in the last 70-year window than in the first 70-year window (note that this varies as we randomly draw the samples; see the results for 100 realizations in Supplementary Fig. 1a).

Averaging over an increasing number of time series. For Fig. 2b, we create 30 time series as the sum of individual white noise time series (mean = 0, s.d. = 1), mimicking measurement- or sampling-induced noise, and the same power-law noise time series (s.d. = 0.5), mimicking the common signal recorded by all the time series. To mimic the decreasing number of observations going further back in time, each time series has a random length but ends at the most recent time point (1,000). We also add one time series of full length (1,000 time units) for the final time series to cover the full period of the original time series. We then take the mean of these time series for which more time series are available with increasing time. See the results for 100 realizations in Supplementary Fig. 1b.

Non-stationary response. Figure 2c reflects a non-stationary or nonlinear response where a variable is observed only when its value is above a certain value (threshold). In this example, when the variable is below 0, no signal is observed. We simulate the variable as the sum of the power-law noise time series (s.d. = 1.2) and white noise (s.d. = 0.5) around a constant mean of −1 units for the first half of the time series, adding a positive linear trend (starting from −1 and ending at 4) for the second half of the time series. This example could represent an ice core where melt rate (the variable) responds to temperature in a nonlinear way; that is, melt is observed only when temperature exceeds zero. Accordingly, when an increasing fraction of the variable rises above zero (due to the positive linear trend), a higher portion of the signal is recorded, and this leads to changes in the measured statistical properties. See the results for 100 realizations in Supplementary Fig. 1c.

Bayesian framework and toy example

First, we derive the Bayes equation as we use it in the section on the Bayesian framework from the standard form of Bayes' theorem.

Bayes' theorem states:

$$P(H|D) = \frac{P(D|H)\,P(H)}{P(D)} \tag{2}$$

where H is the hypothesis and D is the data. We can determine $P(D)$ only by also considering the alternative hypothesis/hypotheses that could also lead(s) to the observed data, $\neg H$:

$$P(D) = P(D|H)\,P(H) + P(D|\neg H)\,P(\neg H) \tag{3}$$

We can divide equation (2) by the same expression with H being replaced by $\neg H$, which eliminates $P(D)$ and yields the odds version of Bayes' theorem (for example, ref. 82):

$$\frac{P(H|D)}{P(\neg H|D)} = \frac{P(D|H)}{P(D|\neg H)} \times \frac{P(H)}{P(\neg H)} \tag{4}$$

There the evidence from the data is included via the likelihood ratio, $P(D|H)/P(D|\neg H)$, which multiplied by the prior odds, $P(H)/P(\neg H)$, provides the posterior odds, $P(H|D)/P(\neg H|D)$.

The alternative hypothesis should include all the plausible alternative hypotheses; that is, $\neg H$ is the union of all the plausible hypotheses or events, $\neg H_1, \neg H_2, \ldots, \neg H_i$. Thus, all corresponding hypotheses have to be enumerated for calculating the term $P(D|\neg H)\,P(\neg H)$:

$$\sum_i P(D|\neg H_i)\,P(\neg H_i) \tag{5}$$

One can convert the posterior odds, PO, to probability, because $P(H|D) + P(\neg H|D) = 1$:

$$P(H|D) = \frac{\mathrm{PO}}{1 + \mathrm{PO}} \tag{6}$$

In our case, H = TP, $\neg H$ = non-TP, and $D = \mathrm{EWS}^+$.

To illustrate how the Bayesian framework can be used for tipping-point detection, we provide an example on how to apply it in Fig. 4 using the AMOC box model as utilized for Fig. 1 and based on ref. 17. In this example, we arbitrarily assume that a study obtained a Kendall's tau of 0.3 from data. This would be considered to be statistically significant ($P < 0.05$) from comparing it with the distribution of Kendall's tau from 1,000 realizations of Fourier surrogates of the corresponding time series. We employ two model configurations (Fig. 1) to simulate AMOC evolution, including with an approaching tipping point (TP model) and without an approaching tipping point but instead a gradual, smooth transition (non-TP model; see Methods, 'AMOC model', for details). Evolution of AMOC strength is represented by the salinity difference between the two boxes in the model. We use the two model configurations to simulate 1,000 time series of AMOC strength for each configuration. For each time series, we calculate the AC1 in running windows (width = 70 data points) and corresponding Kendall's tau, which indicates whether there is an increase in autocorrelation with time, that is, would be interpreted as detection of an EWS. We estimate probability density functions from both samples of Kendall's taus. To estimate the likelihood ratio, we use the probability densities for the Kendall's tau of 0.3 based on the TP and non-TP hypotheses, which would be 0.53 and 2.77, respectively. These would translate into probabilities of 0.1606 and 0.8394 (0.1606 = 0.53/(0.53 + 2.77)), respectively, that is, probability of 0.1606 to observe the EWSs under the TP hypothesis. The ratio of those yields the likelihood ratio; for the assumed Kendall's tau of 0.3 from data (dashed black line in Fig. 4), it is 0.191. This would correspond to moderate evidence for the TP model being more likely to explain the data[107]. However, when the prior odds are not 1, but instead resemble the notion that tipping events are low-probability scenarios, the posterior odds will decrease. For example, if the scenario without tipping point is twice as likely as the TP scenario (for example, when P(TP) is 0.33 and P(non-TP) is 0.66), that is, prior odds of 0.5, then the resulting posterior odds (PO) for an approaching tipping point would be PO = 0.0955. This would correspond to only weak evidence for the TP model being more likely to explain the data[107]. If the hypotheses are mutually exclusive and exhaustive, the resulting probability $P(\mathrm{TP}|\mathrm{EWS}^+) = \mathrm{PO}/(1 + \mathrm{PO}) = 0.087$.

This toy example is only for illustration purpose as we arbitrarily select a Kendall's tau of 0.3. More important, we use one AMOC model where we use constant parameter values to represent TP and non-TP scenarios. In practice, more work needs to be conducted for the choice of appropriate models to estimate the likelihood ratio, and the full parameter space should be sampled. Furthermore, for this illustration we use pre-transition time series with a constant length from the model simulations to calculate the Kendall's tau and the likelihood ratio. Thus, depending on the time it takes for the transition to occur in the model simulations, the Kendall's tau also varies due to varying time to the transition. This means that the distributions shown in the figure are shifted slightly to lower Kendall's tau values. However, as we choose an arbitrary Kendall's tau, there cannot be any inferences made from the example calculation.

Code availability

All code is publicly available via Zenodo at https://doi.org/10.5281/zenodo.14185461 (ref. 108).

References

106. Huybers, P. & Curry, W. Links between annual, Milankovitch and continuum temperature variability. *Nature* **441**, 329–332 (2006).
107. Van Doorn, J. et al. The JASP guidelines for conducting and reporting a Bayesian analysis. *Psychon. Bull. Rev.* **28**, 813–826 (2021).
108. Skiba, V., Hébert, R. & Weinans, E. Code to produce figures for the manuscript 'Ambiguity of statistical early warning signals for climate tipping points'. *Zenodo* https://doi.org/10.5281/zenodo.14185461 (2024).

Acknowledgements

The research of M.R. is supported by the European Research Council (ERC-Synergy project RESILIENCE, proposal no. 101071417) and by the Dutch Research Council (NWO 'Resilience in complex systems through adaptive spatial pattern formation', project no. OCENW.M20.169). This work was conducted as part of the EMBRACER programme, the Earth System Feedback Research Centre, and was financially supported by the SUMMIT programme of the Dutch Research Council (NWO). The research of T.L. is supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (ERC-Starting project SPACE, grant agreement no. 716092) and through the Cluster of Excellence, The Ocean Floor—Earth's Uncharted Interface funded by the German Research Foundation (DFG; EXC 2077, grant no. 390741603). The research of V.S. was supported by German Federal Ministry of Education and Research (BMBF) as Research for Sustainability initiative (FONA); www.fona.de through the Palmod project (FKZ: 01LP2310B). We thank B. Grusdt, A. Dolman, T. Shepherd and P. Zaspel for fruitful discussions. S. Bathiany, A. van der Kaaden and A. Staal are acknowledged for critically reviewing earlier drafts of this paper.

Author contributions

M.R. and T.L. conceived of the study. M.R. and V.S. wrote the first draft of the paper. T.L. and V.S. provided the drafts of the statistical concepts. M.R., V.S., E.W., R.H. and T.L. reviewed and edited the text. V.S., E.W. and R.H. provided the figures.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41558-025-02328-8.

Correspondence should be addressed to Max Rietkerk or Vanessa Skiba.

Peer review information *Nature Climate Change* thanks Fabian Dablander, Juan Rocha and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.