

Anticipating Critical Transitions in Psychological Systems using Early Warning Signals: Theoretical and Practical Considerations

Fabian Dablender¹, Anton Pichler^{2,6,7}, Arta Cika⁴, and Andrea Bacilieri^{2,3,5}

¹*Department of Psychological Methods, University of Amsterdam*

²*Institute for New Economic Thinking at the Oxford Martin School, University of Oxford*

³*School of Geography and the Environment, University of Oxford*

⁴*Department of Engineering Science, University of Oxford*

⁵*Smith School of Enterprise and Environment, University of Oxford*

⁶*Mathematical Institute, University of Oxford*

⁷*Complexity Science Hub Vienna*

This manuscript has been accepted for publication in *Psychological Methods*.

Abstract

Many real-world systems can exhibit tipping points and multiple stable states, creating the potential for sudden and difficult to reverse transitions into a less desirable regime. The theory of dynamical systems points to the existence of generic early warning signals that may precede these so-called critical transitions. Recently, psychologists have begun to conceptualize mental disorders such as depression as an alternative stable state, and suggested that early warning signals based on the phenomenon of critical slowing down might be useful for predicting transitions into depression or other psychiatric disorders. Harnessing the potential of early warning signals requires us to understand their limitations as well as the factors influencing their performance in practice. In this paper, we (a) review limitations of early warning signals based on critical slowing down to better understand when they can and cannot occur, and (b) study the conditions under which early warning signals may anticipate critical transitions in online-monitoring settings by simulating from a bistable dynamical system, varying crucial features such as sampling frequency, noise intensity, and speed of approaching the tipping point. We find that, in sharp contrast to their reputation of being generic or model-agnostic, whether early warning signals occur or not strongly depends on the specifics of the system. We also find that they are very sensitive to noise, potentially limiting their utility in practical applications. We discuss the implications of our findings and provide suggestions and recommendations for future research.

1 Introduction

Inspired by a dynamical systems perspective, researchers in psychology and psychiatry have begun to conceptualize mental disorders such as depression as an alternative stable state (e.g., Borsboom, 2017; Cramer et al., 2016; Hofmann et al., 2016; Kalisch et al., 2019), noting that transitions from a “healthy” into an “unhealthy” stable state may not always be smooth, but can be sudden (e.g., Hayes & Andrews, 2020; Hayes et al., 2007; Helmich et al., 2020; Hosenfeld et al., 2015; Nelson et al., 2017). Such sudden transitions are not only notoriously hard to predict, but can also be hard to reverse. This makes tools that can help us anticipate, and avert, these so-called *critical transitions* highly desirable.

There exist indicators — known as *early warning signals* — that, at least for some systems, can occur prior to such critical transitions. The most widely used early warning signals are based on *critical slowing down*, the phenomenon that some systems return more slowly to their stable state after an external perturbation (Scheffer et al., 2009; Wissel, 1984). Early warning signals based on critical slowing down have been observed prior to transitions in a wide range of systems, for example preceding algal bloom in lakes (Wilkinson et al., 2018), preceding population extinction (Dai et al., 2013; Dai et al., 2012), preceding transitions in the climate (Dakos et al., 2008; Lenton, 2011), and preceding the resurgence of infectious diseases (Harris et al., 2020). This has inspired work in

psychology and psychiatry, where early warning signals have been investigated in the context of major depressive disorder (Kuranova et al., 2020; Schreuder et al., 2020; van de Leemput et al., 2014; Wichers et al., 2016; Wichers et al., 2020), bipolar disorder (Bayani et al., 2017; Curtiss et al., 2019), and sudden gains and losses in psychotherapy (Olthof et al., 2019). Such work is promising and exciting, and several researchers have suggested that early warning signals could be very useful for personalized early intervention in the context of psychiatric disorders (e.g., Hayes et al., 2019; Hofmann et al., 2016; Nelson et al., 2017; Olthof et al., 2019; van de Leemput et al., 2014; Wichers et al., 2019; Wichers et al., 2020).

While early warning signals based on critical slowing down have been suggested as a tool to anticipate critical transitions in a wide range of systems, over the last decade or so, their limitations have become clearer. These limitations help us understand when we can expect early warning signals to occur or not to occur, and are therefore important for practical applications. After explaining the theory behind critical slowing down in Section 2, we provide an overview of the limitations of early warning signal based on critical slowing down in Section 3. Similarly, while early warning signals have been suggested as a tool to monitor in real-time whether patients are about to transition into an unhealthy state, the conditions under which such an approach is feasible in practice remain to be evaluated. In Section 4, we use a simulation study to investigate how factors such as sampling frequency, noise intensity, and the time it takes the system to approach the tipping point may influence the performance of early warning signals in such real-time monitoring situations. Our theoretical and practical investigations have a number of implications for the study and application of early warning signals in psychology and psychiatry. We discuss these implications and provide recommendations for future research in Section 5.

2 Theory of Critical Slowing Down

In the following two sections, we will explain the theory behind critical slowing down using a unidimensional and a multidimensional system as an example. For an extended introduction, we refer the reader to Dablander (2020).

2.1 Unidimensional Systems

We illustrate the theory of critical slowing down using a simple example: modeling the growth of a population of prey under different predation rates (e.g., May, 1977). A key feature of this and many other models is that they can exhibit multiple equilibria, corresponding in our case to multiple population sizes to which the prey can converge over time, given a particular initial population size and a particular value for the predation rate. Such equilibria can either be *stable* or *unstable*, and we refer to these equilibria also as stable or unstable states, respectively. If an equilibrium is stable, the system returns to it after small perturbations. For example, after adding a few more animals to the population of prey, others die due to the limited amount of resources, resulting in a population size that remains constant. If an equilibrium is unstable, on the other hand, the system does not return to the equilibrium after a small perturbation. For example, a prey population of size zero ($x = 0$) is unstable because adding animals of the opposite sex leads to offspring and thus a growing population. The central panel in Figure 1 shows stable (solid lines) and unstable (dashed lines) equilibria, which we denote as x^* , for different predation rates. The grey arrows in the panel illustrate that stable equilibria are attracting, while unstable equilibria are repelling. In the context of our simple example, low predation rates may correspond to a small population of predators, while high predation rates may correspond to a large population of predators. The central panel in Figure 1 reflects the intuition that the stable population of prey should be larger when the predation rate is low compared to when the predation rate is high. For a particular range of predation rates between these two extremes — indicated by the grey shaded region — the system exhibits two stable population sizes; depending on the initial prey population size, the system either converges to one or the other stable state. A necessary condition for multiple stable states and critical transitions are strong reinforcing feedback loops (Kéfi et al., 2016). In our example, individual prey can help each other fend off predators more effectively after reaching a certain population size. This leads to stronger population growth, which in turn boosts their cooperative defense strategy, which leads to stronger population growth etc. Below this population threshold, which is given by the size of the unstable equilibrium, the defense strategy ceases to be

effective, and the prey population collapses. Such a positive relationship between population size and fitness is known as an Allee effect (e.g., Kramer et al., 2009; Stephens & Sutherland, 1999).

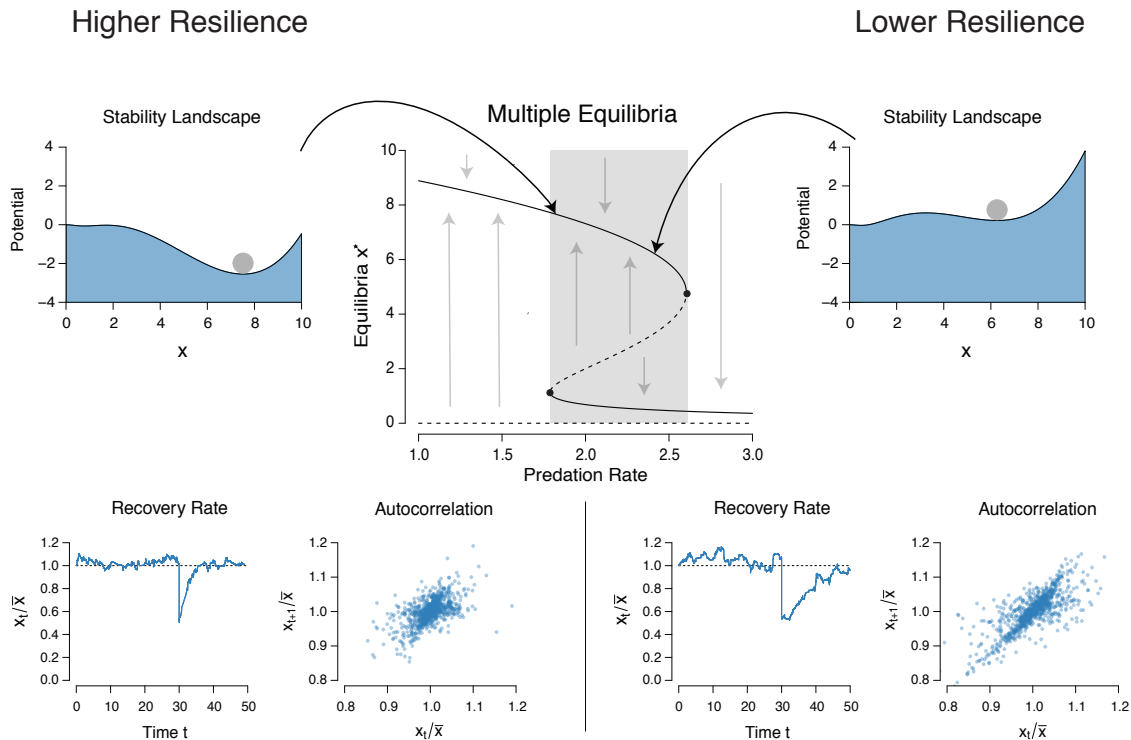


Figure 1: The central panel illustrates the change in stable and hence attracting (solid lines) and unstable and hence repelling (dashed lines) equilibria as the predation rate changes. The top left and right panels show the stability landscape of the system for different predation rates, while the bottom left and right panels show how quickly the system recovers after an intervention, as well as the autocorrelation of the state variable.

Some systems can abruptly change their stable equilibrium as an underlying parameter changes. We show this here for our system, which changes its stable population size as the predation rate changes. The central panel in Figure 1 illustrates this: as we move on the x -axis from a small predation rate to a larger predation rate, we find that around a value of 2.60 — indicated by a black dot — the stable equilibrium $x^* = 4.90$ vanishes (the solid black line above ends) and the system changes to its other stable equilibrium $x^* = 0.44$ (indicated by the black line below). The population of prey is thus driven close to extinction by this miniscule change in the predation rate. A situation in which a small change in an underlying parameter can result in a qualitative change in the stability landscape is known as a *bifurcation*; our specific case of a stable equilibrium vanishing is known as a *saddle-node bifurcation* (Strogatz, 2015). Once the stable equilibrium vanishes, the system moves towards its new stable equilibrium, as the grey arrows in Figure 3 indicate. This is frequently referred to as a *critical transition* (e.g., Scheffer, 2009), and the bifurcation point as a *tipping point* (e.g., Milkoreit et al., 2018; van Nes et al., 2016). How long it takes the system to reach its other stable equilibrium — and thus how *sudden* the change appears to us — depends on the system. For example, after crossing a tipping point a population of yeast can collapse within days (Dai et al., 2012), while the total melting of the Greenland ice sheet would take millennia (Robinson et al., 2012). As the panel illustrates, however, a defining feature of a critical transition is that it is hard to reverse: it is not enough to reduce the predation rate slightly; instead, it has to be reduced to 1.70, at which point another saddle-node bifurcation occurs, leading to a sudden *increase* in population size.

Crossing such tipping points can have catastrophic effects in real-world systems, and indicators that could help us understand whether the system is close to a tipping point would be of great practical importance. The *resilience* of a system is defined as the magnitude of the perturbation the system can withstand without tipping into another equilibrium (Holling, 1973; Scheffer et al., 2015). The top panels in Figure 1 contrast a system that is in a stable state with higher resilience

(left) with one that is in a stable state with lower resilience (right) by means of so-called “ball-in-a-cup” diagrams. Such diagrams visualize the *potential* of a system — colloquially its stability landscape — and a ball (e.g., Strogatz, 2015, pp. 30). The ball represents the state of the system, which moves towards a stable equilibrium with a speed proportional to the local steepness of the landscape. The panel on the top left shows that a perturbation would need to be quite strong to push the high resilient system from its stable state $x^* = 7.52$ to the other stable state at $x^* = 0.78$. In contrast, the “ball-in-a-cup” diagram on the top right illustrates that a smaller push is required for the low resilient system to go from $x^* = 6.26$ to $x^* = 0.49$. In practice, the less resilient system is closer to a tipping point and may need to be nurtured into a more resilient state. Unless we have a realistic mathematical model through which we can assess the effect of perturbations of varying strength, using resilience as an indicator of how likely a critical transition is to occur is difficult. This is because, using our example as illustration, assessing how many animals we could kill until the population collapses by actually killing them would defeat the purpose of preventing such a collapse in the first place.

The *stability* of a system provides a more practical way to assess whether the system is close to a tipping point. Stability is defined as the time the system takes to return to its equilibrium after a small external perturbation (e.g., van Nes & Scheffer, 2007). From the stability landscapes in Figure 1 we see that the more resilient system would return more quickly to its equilibrium state after a small external perturbation than the less resilient system — this is because of the steeper slope of the landscape around the stable equilibrium. The bottom left panels in Figure 1 illustrate this. They show a simulated time-series subject to noise for a system with higher resilience (left) and lower resilience (right). We intervene in the system at time point $t = 30$ and half the prey population. As can be seen, the more resilient system recovers swiftly (at about $t = 35$ the full population size is established again), while the system with lower resilience takes longer to recover. The phenomenon that as a system becomes less resilient — that is, as it comes closer to the tipping point — its return to equilibrium is slower is known as *critical slowing down* (Wissel, 1984).

To quantify the resilience of a system using the concept of stability requires that the system goes out of equilibrium. One way to achieve this is by way of intervention, yet interventions are not always possible or practical in real-world systems. Fortunately, real systems are always subject to countless small external perturbations. We can use these external perturbations as a way to quantify the resilience of a system and find signatures of critical slowing down in empirical time-series. The bottom right panels in Figure 1 illustrate this as an increase in autocorrelation for the system with lower resilience: critical slowing down leads to system states becoming more similar to each other, which is reflected as increased autocorrelation. Similarly, because the system with lower resilience returns more slowly to equilibrium, external perturbations can accumulate and push the system further away from equilibrium, resulting in an increase in variance that can be seen in the bottom panels in Figure 1 before the intervention takes place. Appendix A gives a rigorous mathematical treatment of critical slowing down in unidimensional systems.

2.2 Multidimensional Systems

In contrast to the example above, most real-world systems consist of multiple variables. Mental disorders such as depression, for example, are comprised of various variables related to, among other things, sleep, emotion, and cognition (e.g., Chevance et al., 2020). The dynamics of emotions are viewed as playing a central role in the development of psychopathology (e.g., Trull et al., 2015; Wichers et al., 2015), and we focus on them here. In particular, emotions can have a positive valence, such as cheerful and content, or a negative valence, such as anxious and sad, and we might believe that there exists a tipping point in the mood system such that, when crossed, a person abruptly moves from a stable positive mood state into a stable negative mood state. To formalize such a simplified mood system, we use a four-dimensional version of the Generalized Lotka-Volterra model (e.g., van Nes & Scheffer, 2004) as our toy model:

$$\frac{dx_i}{dt} = 1 + r_i(t)x_i + \sum_{j=1}^4 C_{ij}x_i x_j + \varepsilon_i \quad , \quad (1)$$

where we interpret the variables x_1 and x_2 as cheerful and content, and the variables x_3 and x_4 as anxious and sad (see also Haslbeck & Ryan, 2021; van de Leemput et al., 2014). Equation (1) describes what the rate of change of each variable x_i depends on. $\varepsilon_i \sim \mathcal{N}(0, \sigma_\varepsilon)$ is uncorre-

lated Gaussian noise, and the constant term 1 ensures that the variables are positive with high probability.¹ The matrix C specifies the coupling between the variables. We set

$$C = \begin{pmatrix} -0.2 & 0.04 & -0.2 & -0.2 \\ 0.04 & -0.2 & -0.2 & -0.2 \\ -0.2 & -0.2 & -0.2 & 0.04 \\ -0.2 & -0.2 & 0.04 & -0.2 \end{pmatrix}, \quad (2)$$

which specifies that the positive mood variables x_1 and x_2 reinforce each other ($C_{12} = C_{21} = 0.04$) while they suppress the negative mood variables x_3 and x_4 ($C_{13} = C_{14} = C_{23} = C_{24} = -0.2$), and vice versa. The diagonal of C encodes how strongly x_i^2 influences the rate of change of x_i . All diagonal values are negative ($C_{11} = C_{22} = C_{33} = C_{44} = -0.2$), implying a self-dampening effect for each mood variable. The value of $r_i(t)$ encodes how strongly x_i influences its rate of change. We set $r_1 = r_2 = 1$ so that the self-reinforcing effect of the positive emotions is the same and constant across time, while we set $r(t) \equiv r_3(t) = r_4(t)$ to allow the self-reinforcing effect to change over time for the negative emotions x_3 and x_4 . A lower value for $r(t)$ implies that external perturbations that increase negative emotions are absorbed quickly, while a higher value for $r(t)$ implies that such external perturbations have a more durable effect due to the larger self-reinforcing effect.

Figure 2 shows that r is the key parameter influencing which mood states are stable equilibria. In particular, the left panel shows that, for small values of r , the system is in a state where positive emotions are high (solid red line) and negative emotions are low (solid blue line). This makes sense, because for a small r the self-reinforcing effect of negative emotions is low. As this self-reinforcing effect becomes stronger, however, the system reaches a tipping point at $r = 1.20$ — it abruptly changes into a state in which positive emotions are low and negative emotions are high. The right panel shows this critical transition: an example (scaled) time-series is simulated from the model where $r(t)$ changes (solid black line) starting at day 50. On day 70, the system reaches $r = 1.20$ and abruptly changes from a dominant positive mood state into a dominant negative mood state. Similar to the example in Section 2.1, this change is again hard to reverse: it is not enough to reduce r slightly; instead, it has to be reduced to $r = 0.90$, at which point another saddle-node bifurcation occurs, leading to a sudden change from a dominant negative mood state to a dominant positive mood state.

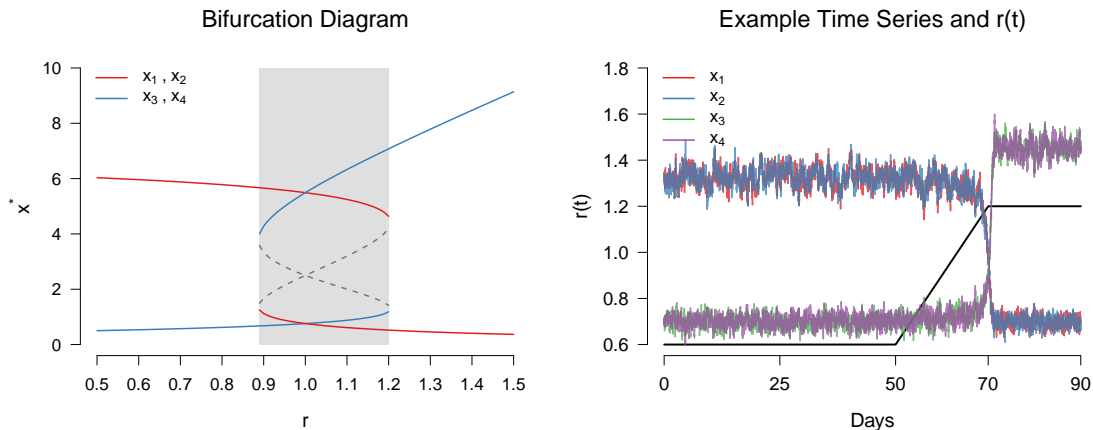


Figure 2: Left: The bifurcation diagram shows how the stable states of the system change as a function of the self-reinforcing effect r of the negative emotions. Right: Time-evolution of r and scaled time-series of the Generalized Lotka-Volterra model given in Equation (1) with $\sigma_\varepsilon = 4$.

As in the unidimensional model described in the previous section, critical slowing down also occurs in this multidimensional model, as illustrated in Figure 3. The top panels show time-series data from a system with higher resilience (left) and the autocorrelation of two of its variables (middle), while the bottom panels show the same plots for a system with lower resilience, resulting

¹The constant term also influences the extent to which the system can exhibit bistability (as highlighted as the grey shaded region in Figure 2); larger values correspond to smaller sets of r values for which the system exhibits two stable equilibria.

in higher variance (left) and higher autocorrelation (middle). In addition to these univariate indicators, a number of multivariate indicators that are sensitive to critical slowing down have been proposed (e.g., Dakos et al., 2010; Kéfi et al., 2014). One such multivariate indicator are the cross-correlations between state variables, which are more pronounced in a system that is close to a tipping point compared to a system that is far away from the tipping point, as illustrated in the right panels in Figure 3. Appendix B gives a rigorous mathematical treatment of critical slowing down in multidimensional systems.

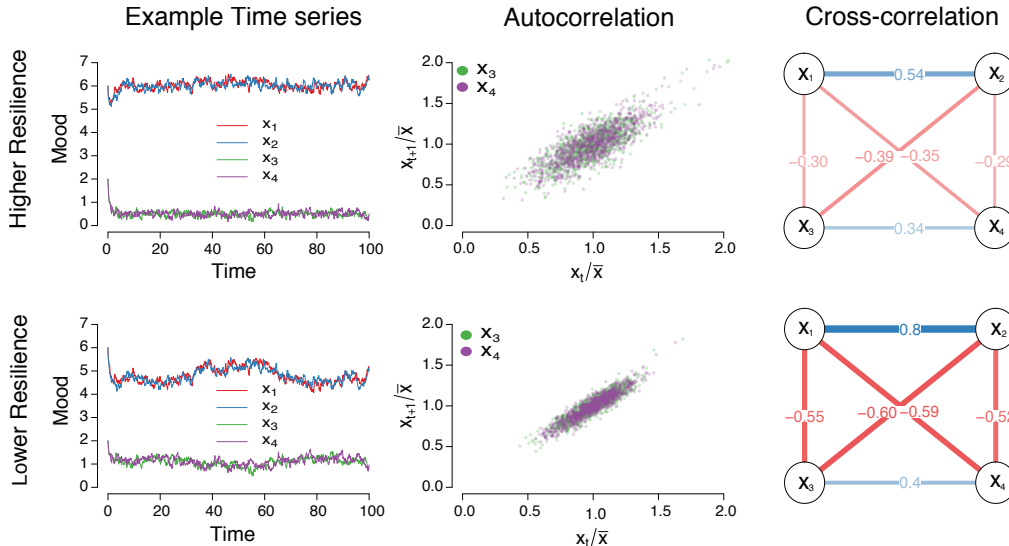


Figure 3: Top: Shows simulated time-series data from our simplified mood model in a state of higher resilience ($r = 0.50$, left), the autocorrelation of two of its state variables (middle), and the cross-correlation of all its state variables (right). Bottom: Shows this for a system with lower resilience ($r = 1.18$). Lower resilience is associated with an increase in variance, autocorrelation, and cross-correlations.

To sum up, critical slowing down is an extremely powerful idea. It suggests that, regardless of the underlying dynamics, as long as the system exhibits multiple stable equilibria and critical transitions between them, we may observe signatures of critical slowing down in time-series data that occur prior to such critical transitions (e.g., Clements & Ozgul, 2018; Scheffer et al., 2012), or allow us to rank the resilience of systems (e.g., Gijzel et al., 2019; Scheffer et al., 2018). Such signatures are known as *early warning signals*, which for the reasons above are sometimes called “generic” in the sense that all we need is time-series data on which to compute simple statistics such as autocorrelation, variance, or cross-correlations. Over the last decade or so, however, the generality of early warning signals based on critical slowing down has been reassessed, and their limitations have become clearer. These limitations have important implications for whether, and when, early warning signals can be useful in practice, and we turn to them in the next section.

3 Limitations of Early Warning Signals based on Critical Slowing Down

In this section, we discuss limitations of early warning signals based on critical slowing down. The key observation is that early warning signals can occur in systems that do not exhibit critical transitions, and that it can fail to occur in systems that do; Table 1 provides an overview. We discuss these two issues in the next two sections.

3.1 Early Warning Signals without Critical Transitions

The systems we have studied in Section 2 exhibit a saddle-node bifurcation as a key underlying parameter — often called *driver* — changes. This type of bifurcation is the most extensively studied case in which critical slowing down precedes a critical transition (Boettiger et al., 2013;

EWS without Critical Transitions	Critical Transitions without EWS
EWS can occur prior to smooth transitions between stable states (Drake & Griffen, 2010; Kéfi et al., 2013).	Strong external perturbations can lead to transitions without EWS (Ditlevsen & Johnsen, 2010; van Nes et al., 2016).
EWS can occur when there is no transition (e.g., Wagner & Eisenman, 2015).	EWS may not occur prior to critical transitions in systems with non-smooth potentials (Hastings & Wysham, 2010).
	Not all variables in a system generally express EWS equally strongly or at all (Boerlijst et al., 2013; Patterson et al., 2021).
	EWS may not occur under correlated or extrinsic noise (Dakos et al., 2012b; O'Regan & Burton, 2018; Qin & Tang, 2018).

Table 1: Summarizes results showing that early warning signals (EWS) can occur prior to non-critical transitions (left) or can fail to occur prior to critical transitions (right).

Scheffer et al., 2009). As shown in Figure 3, as the system approaches the saddle-node bifurcation it takes longer to recover from small perturbations, resulting in critical slowing down which gives rise to early warning signals in the form of increased autocorrelation, variance, and cross-correlations. This is in contrast to the state variables, which barely change as the system approaches the tipping point, as illustrated in Figure 2. Without early warning indicators, it would be difficult to see the critical transition coming. This makes such indicators potentially very useful in practice, and they have been suggested as a tool to anticipate, for example, the sudden onset of depression (e.g., van de Leemput et al., 2014; Wichers et al., 2020).

Critical slowing down is not specific to the saddle-node bifurcation with its implied critical transition, however. Instead, critical slowing down occurs prior to all so-called *zero-eigenvalue* bifurcations (Kéfi et al., 2013; Strogatz, 2015). The *transcritical* bifurcation is one such type of bifurcation: in contrast to the saddle-node bifurcation, it results in a smooth transition between stable equilibria, yet shows critical slowing down (Chisholm & Filotas, 2009; Drake & Griffen, 2010; Kéfi et al., 2013). While exciting recent work tries to identify the type of bifurcation (Bury et al., 2020; Bury et al., 2021), increases in commonly used early warning indicators do not necessarily imply an impending critical transition. In fact, Kéfi et al. (2013) showed that the system need not even experience a bifurcation at all, but that a smooth, nonlinear change in the stable equilibrium due to a change in an underlying parameter can suffice to elicit critical slowing down.

An increase in early warning indicators such as autocorrelation and variance can also occur due to a number of other factors. For example, the variance of the system might simply increase due to stronger noise or an increase in the variance in key underlying drivers (see for example Boulton et al., 2013). When this happens, we might be misled into thinking that a critical transition is impending, when in fact it is not. For this reason, it is advised to assess if multiple indicators, such as variance and autocorrelation, increase instead of relying on a single indicator (e.g., Ditlevsen & Johnsen, 2010). However, increases in multiple indicators without critical transitions are still possible, and can thus lead to false alarms (e.g., Boettiger & Hastings, 2012a; Wagner & Eisenman, 2015).

3.2 Critical Transitions without Early Warning Signals

While signalling non-critical transitions or falsely signalling critical transitions can be problematic, failing to signal critical transitions might be catastrophic. An obvious case in which a system experiences a critical transition without critical slowing down is the following: as long as the system exhibits two stable equilibria, there is always the possibility that a strong external perturbation pushes the system from one stable equilibrium into another, independently of any change in the system dynamics (Boettiger & Hastings, 2012a; Ditlevsen & Johnsen, 2010; van Nes et al., 2016). In the context of our mood model, the sudden death of a close relative may push a person from a positive state into a negative state almost immediately and without warning. As shown in Figure 2, the closer we are to the bifurcation point, the smaller such a perturbation — for example corresponding to a less severe negative life event in our context — has to be to push the system into the other stable equilibrium. Since real-world systems are always subject to external perturbations

that may tip the system into an alternative stable equilibrium at any time, one cannot exactly predict the occurrence of critical transitions. While this should not come as a surprise, it deserves repetition; it is the reason why careful authors speak of *anticipating* instead of *predicting* critical transitions (e.g., Scheffer et al., 2012).

There are more subtle cases in which critical transitions can occur without critical slowing down, however. As explained in Section 2, critical slowing down assumes that the system is in equilibrium, with small noise constantly perturbing it after which it may recover. Crucially, it assumes that factors influencing the system change slowly compared to the rate with which the system returns to its equilibrium. This means that the system can respond to small external perturbations in time, that is, return (close) to the stable state before the environment changes — the system “tracks” the continuously changing stable state (Perryman, 2015). This need not be the case, however. If instead the environment changes so rapidly that the system cannot adjust back to the stable state in time, it can undergo *rate-induced* tipping (Ashwin et al., 2012; Scheffer et al., 2008; Siteur et al., 2016; van der Bolt & van Nes, 2021). For example, a critical factor for the stability of certain systems in future climate may not only be the magnitude, but also the rate of global warming (e.g., Lohmann & Ditlevsen, 2021; Luke & Cox, 2011). In the context of our mood example, we may find that a person stays in a positive mood state as stress increases — while staying below a critical level — slowly over days. This increase in stress is slow enough so that the positive emotions can balance out a rise in negative emotions. However, should stress increase within hours, it may be that positive emotions are too slow to balance out a rise in negative emotions, and the person transitions into a negative mood state, even though the stress level always stayed below the critical level. In this case, it is not the absolute level of stress that is critical, but rather its rate of change. In contrast to bifurcation-induced critical transitions, the stability landscape does not change before rate-induced transitions; the concept of resilience as outlined in Section 2 may thus be inadequate for systems where rate-induced transitions are possible (Siteur et al., 2016). As a corollary, rate-induced transitions therefore need not be anticipated by conventional early warning signals (Boulton et al. (2013); Ashwin et al. (2012); but see Ritchie and Sieber (2016); Siteur et al. (2016)).

Another subtlety that has important ramifications for early warning signals based on critical slowing down concerns the relationship between resilience and stability. As we have seen in Section 2.1, critical slowing down is a result of a decrease in a system’s stability as it approaches a tipping point. In our example, the resilience and stability of the system are positively correlated: as stability decreased with increasing predation rate, so did resilience. While our simple predator-prey system is subject to only one slowly changing driver — predation rate — real-world systems are subject to multiple drivers. Driving a yeast population to collapse, Dai et al. (2015) showed that the relation between stability and resilience can change depending on the driver (see their insightful Figure 4), and demonstrated experimentally that the performance of early warning signals based on critical slowing down markedly decreases when stability decreases more slowly relative to resilience. Crucially, they also found that when the system is subject to two drivers that change in opposite directions, stability and resilience can become negatively correlated: resilience can *decrease* while stability actually *increases* as the tipping point is approached (see Dai et al., 2015, Figure 5). This results in decreasing early warning indicators, and may thus incorrectly signal that no tipping point lies ahead. In the context of our psychological example, stress may increase, moving the system towards the tipping point, but social support may increase as well, moving the system away from the tipping point, potentially resulting in a negative correlation between stability and resilience. This important nuance in the relationship between stability and resilience points to the need of understanding the underlying drivers of a system and their interactions.

Early warning signals based on critical slowing down may fail to anticipate critical transitions also for another reason. Critical slowing down requires that the potential — frequently called the stability landscape — of a dynamical system not only changes, but changes smoothly as an underlying driver changes (Hastings & Wysham, 2010). Figure 1 shows such stability landscapes for a system with higher and a system with lower resilience; the assumption of smoothness means that, as we vary the predation rate, there should not be any “breaks” in the stability landscape. Yet non-smooth potentials are possible; they occur, for example, in chaotic systems, that is, in systems whose time-evolution depends sensitively on initial conditions. Hastings and Wysham (2010) argued that chaotic systems are only a small subset of systems that have non-smooth potentials. They illustrated the lack of critical slowing down on three models that likely do not have smooth potentials, but that undergo critical transitions. They argued that a large class of

real-world systems might have non-smooth potentials, and that one therefore cannot assume a priori that critical slowing down precedes critical transitions in real-world systems. Hastings and Wysham (2010) stressed the importance of building mathematical models of the phenomenon one is interested in, as this can suggest whether the system can be adequately described by a model exhibiting a smooth or a non-smooth potential. Titus and Watson (2020) echoed the call by Hastings and Wysham (2010) for the need of modeling after showing that the autocorrelation and variance can *decrease* prior to transitions, a phenomenon they term *critical speeding up*.

Even if we assume that potentials are smooth and critical slowing down exists, not every variable in a multidimensional system might express early warning signals such as an increase in autocorrelation or variance equally strongly — or at all. Boerlijst et al. (2013) studied a three-species model with juvenile prey, adult prey, and a predator that only attacks adult prey. They varied the internal parameter controlling the predator’s death rate such that the predator population undergoes a saddle-node bifurcation and becomes extinct. Strikingly, they did not observe critical slowing down in either the predator or the adult prey population. Instead, critical slowing down could only be observed in the juvenile population, which is arguably the most irrelevant variable, since the predator preys solely on adults.² Motivated by this finding, Patterson et al. (2021) provided a rigorous mathematical analysis of the conditions under which early warning signals are expected to occur. Importantly, they found that the expected strength of early warning signals decreases with the square of the total number of variables comprising the system (not all of which are usually observed). These theoretical insights have important practical implications. Returning to our mood example, in the real world we might observe early warning signals most strongly in the variable “cheerful”, but barely or not at all in the other three variables, all four of which are just a small part of the larger mood system. Without a good understanding of the system that can inform which variables one should monitor, early warning signals will generally be harder to observe in complex, high-dimensional systems.

Another important limitation of critical slowing down is that its occurrence depends on the type of noise. For example, while Boerlijst et al. (2013) observed critical slowing down in the juvenile prey for uncorrelated Gaussian noise added to all three populations, critical slowing down is not observed under noise that is fully correlated. Correlated noise is common in real-world systems, as external perturbations generally do not affect system components independently. For example, suppose it is shown that a model for panic attacks based on a reinforcing feedback loop between autonomic arousal and perceived threat (Robinaugh et al., 2019) exhibits critical transitions and critical slowing down. In practice, however, the system the model is describing is always subject to noise, and because the system is much more complicated than the model — for example, the model does not include all relevant variables — this noise will be correlated. Since correlated noise can markedly suppress early warning signals, this may have considerable practical implications.

Early warning signals are usually derived under the assumption of additive white noise, neglecting the underlying noise mechanism. Taking the noise mechanism into account, one can distinguish between *extrinsic* and *intrinsic* noise (e.g. Boettiger, 2018). Extrinsic noise encodes changes in the environment; it is thus shared by all system components, inducing correlations between them (Qin & Tang, 2018). In our simple model illustrated in Figure 1, extrinsic noise would correspond to perturbations of the predation rate. Intrinsic noise, on the other hand, is due to randomness in the system itself, for example due to stochastic births and deaths in the population (O’Regan & Burton, 2018). Qin and Tang (2018) found that early warning indicators can fail to anticipate critical transitions under extrinsic noise and can also lead to false alarms; notably, the situation becomes worse as the size of the system increases (see also Patterson et al., 2021). O’Regan and Burton (2018) further showed that, while autocorrelation is robust to different noise forms, the variance is not — it can increase, decrease, or stay constant before a saddle-node bifurcation, depending on how the system dynamics interact with the dynamics of the noise (see also Dakos et al., 2012b).

The limitations above strongly put into question the notion that early warning signals can be used as a generic tool to anticipate critical transitions in a large class of systems. Instead, we need to build up a sufficient understanding of the system under study in order to assess whether and under what conditions early warning signals are likely to anticipate critical transitions — assuming that critical transitions can in fact occur in the system under study. Yet even if we had a sufficient theoretical understanding that makes the occurrence of early warning signals before critical transitions likely, it is currently unclear how well early warning signals would perform in online monitoring settings in anticipating such transitions. To assess this, we turn to a simulation

²We have replicated their results and extended them to multivariate indicators in Appendix C.

study in the next section.

4 Investigating the Performance of Early Warning Signals

The utility of early warning signals does not only depend on the theoretical considerations outlined in the previous section. After all, detecting a rise in indicators such as autocorrelation or variance is also a statistical challenge. It is therefore important to understand what factors may influence the performance of early warning signals in practice. We address this issue by simulating from the Generalized Lotka-Volterra which we introduced in Section 2.2 as a simplified model for mood. While bistable systems of this type have become popular in theoretical psychology (e.g., Borsboom, 2017; Cramer et al., 2016; Nelson et al., 2017), real psychological systems are more complicated and give rise to much messier data. However, this simplified setup allows us to study the purely statistical challenges of applying early warning signals in practice in a straightforward manner. Previous simulation studies assumed that the system has transitioned and that the time point of the transition is known (e.g., Boettiger & Hastings, 2012b; Clements et al., 2015; Peretti & Munch, 2012). This situation is less interesting in practical applications, however, where we wish to assess whether early warning indicators signal a critical transition before it actually occurs. In our simulation study, we mimic an online monitoring setting where data comes in sequentially and early warning indicators are computed with every new data point. Our goal is to understand the conditions under which such an approach is feasible in practice.

4.1 Simulation Setup

We set up our simulation study to shed light on the following questions. First, what is the respective influence of increasing noise and decreasing sampling frequency on the performance of early warning signals? Since detecting a rise in indicators is a statistical problem, increasing noise and decreasing sampling frequency likely have a detrimental effect, but the extent of their respective negative influence is unclear. Second, how does the extent of baseline data and the length of the transition period influence the performance of early warning signals? Since we mimic an online monitoring setting in which we sequentially collect data and do not assume knowledge about the occurrence of the tipping point, we need to test a potential rise in indicator against the indicator at baseline. This requires a good estimate of the indicator both at baseline as well as during the transition period. We therefore expect that more baseline data as well as a longer transitioning period result in an increased performance of early warning indicators. Finally, what early warning indicators best anticipate a transition for our system? While there are a large number of early warning indicators — with more and more being developed — we assessed the performance of the following widely used early warning indicators: autocorrelation, variance, skewness, and kurtosis as univariate indicators (Carpenter & Brock, 2006; Guttal & Jayaprakash, 2008; Scheffer, 2009); the average absolute value of all cross-correlations (Dakos et al., 2010), the largest eigenvalue of the covariance matrix (Chen et al., 2019), spatial variance, spatial skewness, and spatial kurtosis as multivariate indicators (Guttal & Jayaprakash, 2009; Kéfi et al., 2014); see Appendix D for mathematical definitions. Note that we estimated all univariate indicators on x_1 , and that using any other variable does not markedly change the results (a fact that is specific to our system).

We sampled the system using Euler’s method with a time step of $\Delta t = 0.01$. Table 2 gives an overview of the early warning indicators we used and the parameters we varied in the simulation. We have categorized the parameters into *uncontrollable* parameters that, in real applications, are usually outside of the researcher’s control and *controllable* parameters that are more likely to be under the researcher’s control. For example, while the extent of unwanted variation (i.e., noise) and the time it takes the system to approach the critical transition can (usually) not be influenced by the researcher, design choices regarding the extent of the baseline or the sampling frequency as well as statistical choices such as the rolling window size are more likely to be under the researcher’s control.

We studied different noise intensities $\sigma_\varepsilon \in [4, 6, 8, 10]$, which in practice may partly arise due to different degrees of measurement error. The extent of the observed noise depends on σ_ε and Δt , and so cannot be taken in absolute terms to reflect some real-world process, but must be interpreted in a relative manner.³ In experience sampling studies common to psychopathology research, the

³To give more intuition for the noise intensities, note that the empirical standard deviation σ_{x_1} of x_1 at $r = 0.60$ corresponding to the levels of σ_ε are $[0.34, 0.51, 0.69, 0.87]$. Transforming these into signal-to-noise ratios μ_{x_1}/σ_{x_1} —

Early Warning Indicator	Source
Autocorrelation and Variance	Scheffer et al. (2009)
Skewness and Kurtosis	Guttal and Jayaprakash (2008)
Cross-correlation	Dakos et al. (2010)
Dominant eigenvalue of covariance matrix	Chen et al. (2019)
Spatial-Variance, Spatial-Kurtosis, and Spatial-Skewness	Kéfi et al. (2014)
Parameter	Values
<i>Uncontrollable</i>	
Noise intensity σ_ε	4, 6, 8, 10
Transition Period	10, 25, 50 days
<i>Controllable</i>	
Sampling Frequency	1x, 5x, 10x per day
Baseline	25, 50, 100 days
Rolling Window Size	10, 25, 50 days

Table 2: Shows the early warning indicators we used and the parameters we varied in the simulation study.

window within which we can query a person ranges from 8am to 10pm, that is, there are 15 hours within which we can sample the system. We studied the behaviour of early warning indicators under sampling frequencies of every 90, 180, and 900 minutes. These sampling frequencies, which are standard in experience sampling studies, yielded 10, 5, and 1 observations per day, respectively.

We say that the system is at baseline when $r = 0.60$, and we varied the length of the baseline to be 25, 50, or 100 days. We implemented different transition periods by linearly changing the stress parameter r from $r = 0.60$ at baseline to $r = 1.20$ with varying steepness. We varied this transition period in the simulation to be 10, 25, and 50 days. For a transition period of 10 days and sampling frequencies of 1, 5, or 10 times per day this resulted in 10, 50, or 100 observations, respectively; for a transition period of 25 days we would collect 25, 125, or 250 observations; and for a transition period of 50 days this would yield 50, 250, and 500 observations, respectively. Figure 4 illustrates three examples with different sampling frequency and noise intensity. From left to right, the simulated time-series are generated with decreasing noise intensity and decreasing sampling frequency. In the simulation, we studied all possible combinations of all parameters. Note that with increasing noise, the chances of transitioning before or after the theoretical bifurcation point at $r = 1.20$ increase.

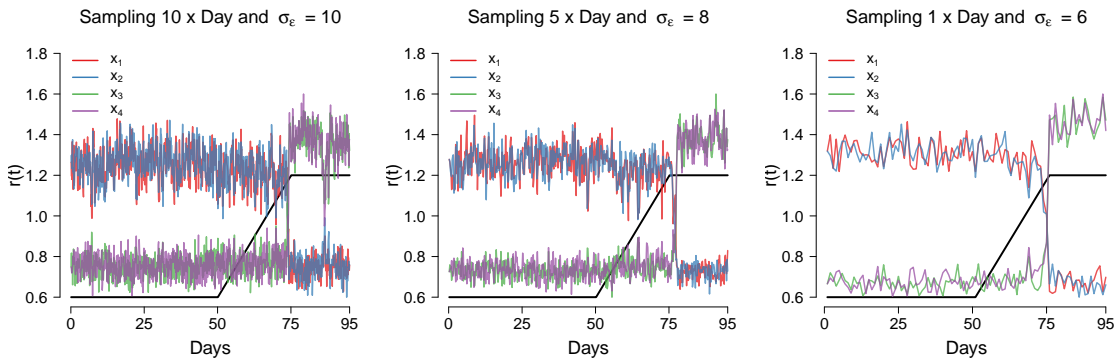


Figure 4: Three example settings in our simulation study. The bifurcation parameter r increases linearly after a baseline period. From left to right, the (scaled) simulated time-series are shown under decreasing noise intensity and decreasing sampling frequency.

While we interpret parameters such as the baseline or the sampling frequency in terms of days with μ_{x_1} being the empirical mean — yields [17.70, 11.60, 8.70, 6.80].

for illustration purposes, there actually is no “correct” time scale in our model. This is because it does not adequately model a real psychological system, and so we cannot use the variables to define an appropriate time scale of the dynamics. For example, if our model were in fact an adequate representation of the mood system and the variable “anxious” were to change on the level of hours, then we could use this knowledge to define what extent of simulated data would correspond to one day. Since this is not the case, however, it is not possible to interpret our results as pertaining to a particular baseline or sampling frequency; instead, the results have to be interpreted relative to each other.

4.2 Statistical Analysis

We used the simplest and most commonly employed method to test for early warning signals, which is based on rolling windows. In particular, for a given rolling window size and for each new data point, we calculated the early warning indicator and signalled a transition when its size exceeds the size of the indicator at baseline for a particular decision threshold. We studied different thresholds, saying that an early warning indicator signalled a critical transition when its current value is $\sigma \in [0.25, 0.50, \dots, 5.75, 6]$ standard deviations above the early warning indicator value at baseline. This allowed us to draw receiver operating characteristic (ROC) curves, illustrating how the true positive and false positive rate of early warning indicators changed as we changed the σ -threshold. We varied the rolling window sizes to be 10, 25, or 50 days, but never more than half the baseline. The rolling window size is an important parameter in the statistical analysis; small window sizes lead to noisier estimates, while large window sizes lead to a reduced sensitivity to detect changes in the indicator.

We analyzed the statistical properties of early warning signals by studying 500 critical transitions from a positive to a negative mood state and 500 cases for which no transition occurs for each parameter combination. As can be seen in Figure 4, the system does not always transition *exactly* at the bifurcation point ($r = 1.20$) due to noise. Therefore, we used a change-point method on the whole time-series to assess the exact time at which the system transitioned (Killick & Eckley, 2014); it is this actual transition that we wish to anticipate, not the noise-free theoretical one. We provide two analyses. First, we used ROC curves to illustrate how the true positive and false positive rate of early warning indicators changed for different conditions as we varied the decision threshold. We refer to the situation where an early warning signal indicated a transition when in fact the system did not transition at any point in time as a false positive. Conversely, we refer to the situation where an early warning signal was followed by a transition as a true positive. Second, in an online monitoring setting it is also important to know how much in advance early warning indicators would signal an impending critical transition. To assess this, we computed for all settings how far in advance an actual transition was signalled.

4.3 Simulation Results

We first give an overall picture of how the performance of early warning indicators varied as a function of the sampling frequency, transition period, and noise intensity. To do so, we combined two univariate and two multivariate indicators into a single indicator. We followed the methodology proposed by Drake and Griffen (2010) for combining indicators and created a new early warning indicator by summing the z -values of the autocorrelation, variance, cross-correlation, and the dominant eigenvalue of the covariance matrix. Figure 5 shows the ROC curves associated with the combined indicator as a function of the theoretical transitioning period — that is, the time between $r = 0.60$ and $r = 1.20$ — the noise intensity σ_ε , and the sampling frequency. An ROC curve summarizes the performance of an indicator for various decision thresholds. The points in the ROC curves shown in Figure 5 denote different σ -thresholds; for each such curve, the leftmost point means that we took an increase of a standard deviation of $\sigma = 6$ in the combined indicator compared to baseline as a signal of a transition; the rightmost point denotes a $\sigma = 0.25$ threshold. The larger the threshold, the more hesitant one is to signal a transition. This results in a trade-off between the false positives rate (shown on the x -axis) and the true positive rate (shown on the y -axis), such that larger σ -thresholds result in fewer false positives but also fewer true positives. The black points indicate the frequently used 2σ -threshold (e.g., Clements et al., 2019; Drake & Griffen, 2010). All results in Figure 5 are with respect to a 100 day baseline and a 50 day rolling window.

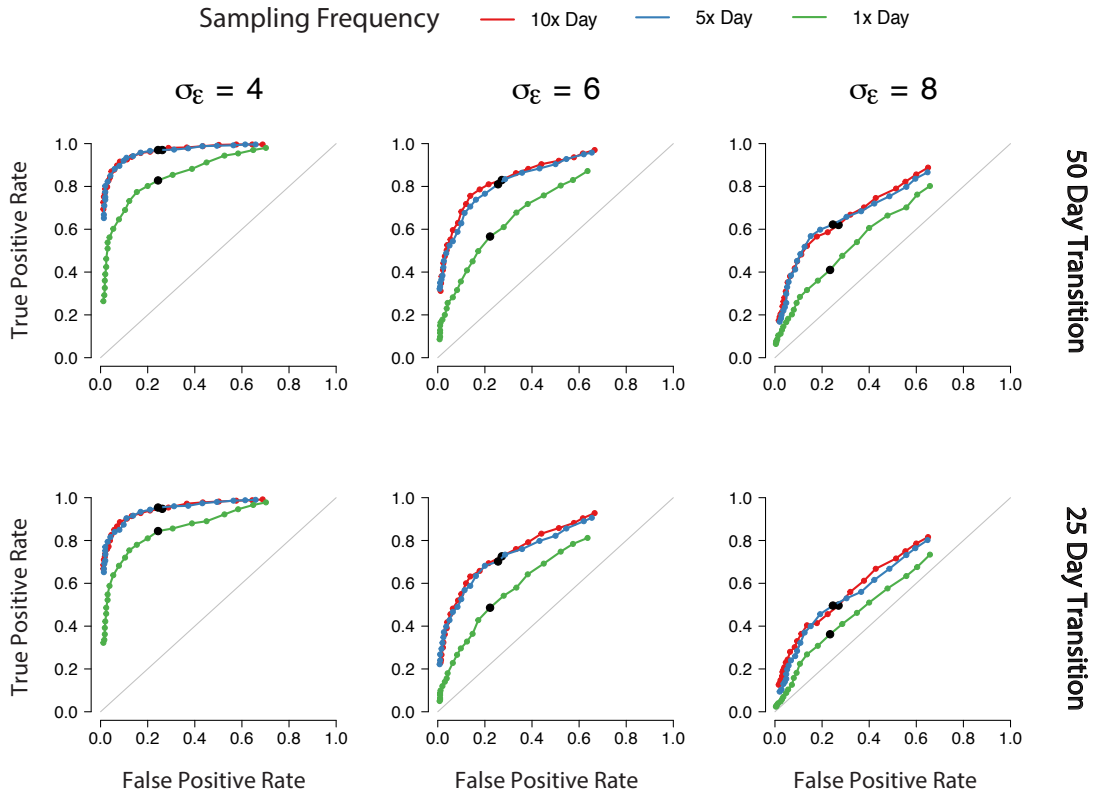


Figure 5: Shows the ROC curves for the combined indicator as a function of the theoretical duration of the transitioning period and $\sigma_\varepsilon \in [4, 6, 8]$. Black dots indicate the 2σ -threshold. All results are with respect to a 50 day rolling window and a 100 day baseline.

As Figure 5 shows, an increase in noise intensity σ_ε resulted in a substantially reduced performance of the combined indicator, regardless of the length of the transitioning period. A decrease in the transitioning period also had a detrimental effect on performance, but this depended on the extent of the noise: for $\sigma_\varepsilon = 4$, there was barely a difference in performance between a transitioning period of 50 and 25 days. However, a shorter transitioning period implied a decreased performance as the noise increased. This is because a shorter transitioning period, implying a steeper slope from baseline to the tipping point, is associated with fewer observations compared to a longer transition period; thus, increased noise can have a more detrimental effect. Interestingly, halving the sampling frequency from 10 times per day to 5 times per day barely reduced performance, while sampling only once per day did so considerably. This suggests that, in practical applications where system-specific features such as the extent of noise and the transition period are fixed, reducing the sampling frequency up to a point may decrease performance only marginally, but that after this point is reached performance may drop substantially.

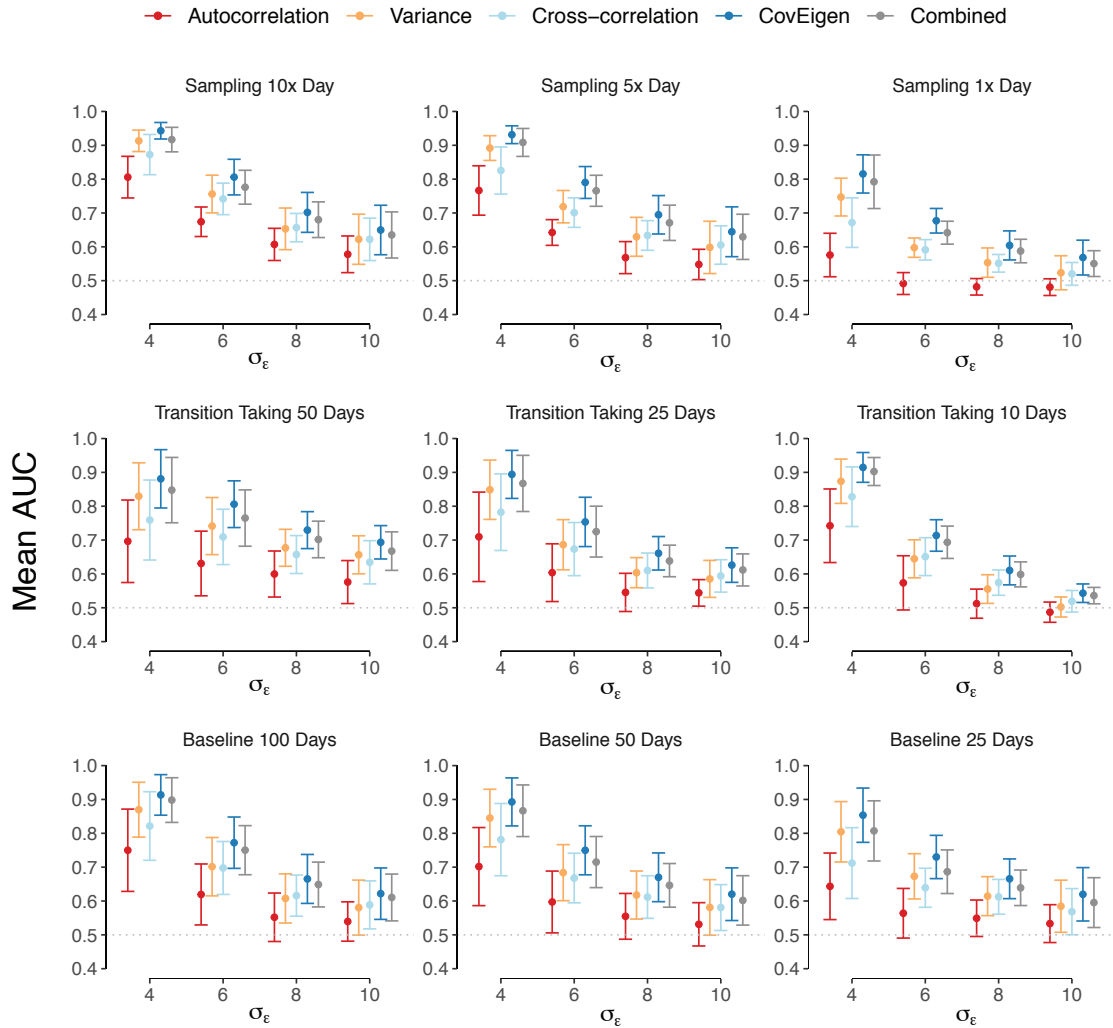


Figure 6: Shows Area under the Curve (AUC) of selected early warning indicators across sampling frequencies (top panels), transitioning periods (middle panels), and baselines (bottom panels). Results are averaged across simulation runs and parameters not depicted in the particular panels, with error bars denoting one standard deviation across these runs. Grey dotted lines indicate chance performance. CovEigen uses the dominant eigenvalue of the covariance matrix as early warning signal.

A widely used performance metric is the *Area Under the Curve* (AUC): an AUC of 0.50 implies chance performance (indicated by the grey diagonal lines in Figure 5), while an AUC of 1 implies perfect detection capabilities. To understand how various early warning signals performed under a range of settings, Figure 6 shows the average AUC for selected early warning indicators for different sampling frequencies (top panels), transitioning periods (middle panels), and baselines (bottom panels) averaged across combinations of all other parameter settings; error bars denote one standard deviation across these configurations, and the grey dotted lines indicate chance performance.⁴ The dominant eigenvalue of the covariance matrix (CovEigen) performed best, closely followed by the combined indicator, the variance, and the cross-correlation, with the autocorrelation trailing. This ordering in indicator performance holds across all settings; see Appendix E for results concerning all other early warning indicators.

The top panels in Figure 6 show that as the sampling frequency decreased, the performance of all early warning indicators decreased. This is also what we observed in Figure 5, and it makes sense: fixing the values of other parameters such as the noise intensity and the transition period, a reduced sampling frequency implies fewer observations and hence a more detrimental impact of noise, thereby reducing performance. Our simulation study mimics real-world settings

⁴Note that because our ROC curves do not touch the (0,0) and (1,1) coordinates, we linearly interpolate by adding them before calculating the AUC.

where certain parameters (e.g., noise intensity, transition period) are not under the control of the researcher, and in such settings it is not a reduced sampling frequency *per se* that leads to worse performance, but the effects a reduced sampling frequency has on the quality of the gathered data. The autocorrelation is the indicator that was least robust to a decrease in sampling frequency, hovering close to chance performance for all levels of noise. The middle panels show that, as the system approached the bifurcation point more quickly, the performance of all early warning indicators decreased as the noise became more pronounced, mirroring the results shown in Figure 5. The bottom panels show that a decrease in baseline resulted in a decreased performance of all early warning indicators.

The most salient result of this analysis is that all early warning indicators suffered considerably from increased noise. This effect seems to be nonlinear, with a substantial drop in performance when increasing the noise level from $\sigma_\varepsilon = 4$ to $\sigma_\varepsilon = 6$, and less pronounced decreases in performance with further increases in noise. From Figures 2 and 4 we see that the extent of the noise did not have a drastic influence on the variability of the state variables; at least not to the extent that one can easily intuit that early warning indicators would perform excellently at $\sigma_\varepsilon = 4$, but would basically be useless at $\sigma_\varepsilon = 10$. Glancing at the variability of the system under $\sigma_\varepsilon = 10$, one may argue that this still represents an ideal case when compared to real-world settings, potentially undermining the usefulness of early warning indicators to anticipate critical transitions in noisy online monitoring settings.

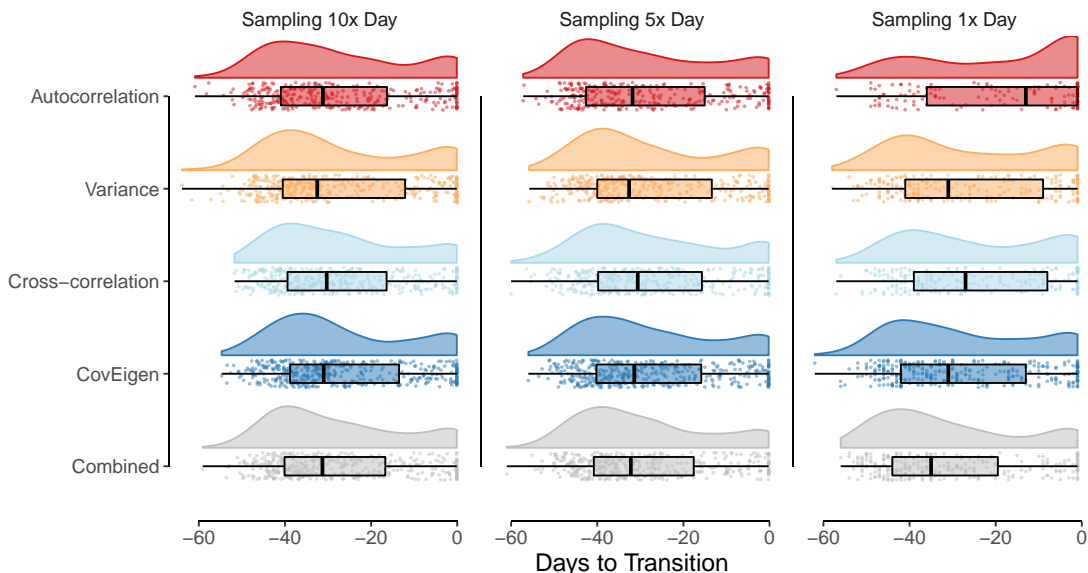


Figure 7: Illustrates when early warning indicators first signalled a transition using a 2σ -threshold, a 50 day rolling window, a 100 day baseline, a theoretical transition period of 50 days, and noise intensity $\sigma_\varepsilon = 6$ for different sampling frequencies. CovEigen uses the dominant eigenvalue of the covariance matrix as early warning signal.

In the analysis above, we have counted as a success whenever an early warning indicator correctly signalled a transition. This neglected how far in advance a transition is anticipated; it clearly matters in practice whether a transition occurs, say, 20 days after an early warning signal or only 2 days after. We now focus on the simulation runs in which an early warning signal occurred prior to the critical transition, that is, we focus on true positives. Figure 7 illustrates how far in advance early warning indicators signalled a critical transition using a 2σ -threshold, a 50 day rolling window, a 100 days baseline, a theoretical transition period of 50 days, and a noise intensity of $\sigma_\varepsilon = 6$. We see that the early warning indicators signalled a transition far in advance of the actual tipping point, and that this pattern was fairly consistent across indicators and sampling frequencies. As the true positive rate decreases with a decrease in the sampling frequency, the number of data points in Figure 7 decreases from the left to the right panel. This is especially the case for the autocorrelation, which showed the largest decrease in the number of true positives and was most likely to signal the transition only shortly before it actually occurred for the lowest sampling

frequency. This pattern is consistent with Figure 6, where the performance of the autocorrelation decreased most strongly as sampling frequency was reduced. Note that the early warning indicators sometimes signalled a transition further ahead (e.g., 60 days) than the theoretical transition period (here, 50 days). As discussed in Section 4.1, noise can make the system transition after (or before) the theoretical tipping point, which — combined with an indicator signalling a transition early — explains the occasionally large advance in anticipation.

While Figure 7 illustrates one particular setting, Table 3 shows how much earlier the combined indicator signalled a transition across all noise intensities, theoretical transitioning periods, and sampling frequencies for a 2σ -threshold. For example, for a theoretical transition period of 50 days, sampling ten times per day, and noise intensity $\sigma_\varepsilon = 4$, the combined indicator signalled a transition about 31.47 days in advance, on average, using a 2σ -threshold. Of course, how far in advance the indicator suggests a transition depends on the decision threshold one uses. For a sufficiently large threshold, more data is needed and early warning signals will indicate a transition closer to when it actually occurs — if at all, since a larger threshold also results in more actual transitions being missed. Here, we focus on the frequently suggested and generally well-performing 2σ -threshold. The key observation is that as the noise increased, the indicator signalling a transition occurred later and later. This pattern holds for all sampling frequencies and transitioning periods, and it makes sense: reduced noise leads to increased statistical power to detect a change in the early warning indicator compared to baseline.

Transition Period	Sampling 10x Day			Sampling 5x Day			Sampling 1x Day		
	10	25	50	10	25	50	10	25	50
$\sigma_\varepsilon = 4$	8.77 (3.81)	17.09 (8.48)	31.47 (16.12)	8.91 (3.7)	17.3 (8.7)	31.81 (16.39)	9.93 (4.01)	18.8 (9.27)	31.7 (18.23)
$\sigma_\varepsilon = 6$	8.36 (4.32)	16.29 (8.76)	27.89 (16.26)	8.48 (4.26)	16.17 (8.91)	28.84 (16.04)	8.69 (4.61)	14.71 (9.81)	26.63 (17.57)
$\sigma_\varepsilon = 8$	6.97 (4.96)	14.19 (8.84)	24.61 (15.59)	7.06 (4.91)	14.57 (8.78)	24.55 (15.7)	6.44 (4.65)	13.08 (9.21)	21.19 (16.27)
$\sigma_\varepsilon = 10$	6.09 (4.43)	11.51 (8.19)	21.11 (14.53)	6.03 (4.54)	11.67 (8.11)	20.96 (14.69)	5.24 (4.2)	10.46 (7.61)	18.92 (15.04)

Table 3: Shows the average difference in days between the actual transition and when the combined indicator first signals a transition using a 2σ -threshold across different noise levels, sampling frequencies, and theoretical transition periods. The average is taken across simulation runs, baselines, and rolling window sizes; the standard deviation across these configurations is shown below in parenthesis.

4.4 Simulation Discussion

To our knowledge, ours is the first simulation study that explicitly studied how well early warning indicators perform when used in an *online* monitoring setting, rather than in settings where the occurrence of the tipping point is known. Using an ecological model and assuming knowledge of a tipping point, Peretti and Munch (2012) also found that the performance of early warning indicators to anticipate critical transitions was substantially reduced under increased noise. Furthermore, they found that a reduction in sampling frequency had a detrimental effect on the performance, and that the performance of the autocorrelation was most strongly affected by a reduction in sampling frequency, which is substantiated by our findings. In a similar spirit, Clements et al. (2015) simulated from the logistic equation with grazing in an ecology context and studied how a reduction in sampling frequency as well as spatial subsampling of the population impacted the performance of early warning indicators, assuming knowledge of the tipping point. They found that a decrease in sampling frequency had a more detrimental effect than spatial subsampling. An encouraging result is reported by Brett et al. (2018), who showed that several early warning indicators were robust to reporting errors and aggregation in anticipating epidemic transitions;

future work may wish to assess the robustness of early warning signals to errors that are more specific to psychology.

In our simulations, the number of measurements taken until the system reached the bifurcation point $r = 1.20$ and the rate with which r changes — the *forcing rate* — were not independent. For example, a transition period of 10 days with a sampling frequency of 10 samples per day resulted in 100 observations from the initial change in r to $r = 1.20$, while a transition period of 25 days with the same sampling frequency resulted in 250 samples. This implies that the rate of change in r was smaller in the latter case than in the former. In real-world settings where we wish to anticipate future transitions, these two factors — the forcing rate and the time it takes the system to reach the tipping point — cannot be disentangled, and so we have not done so here. In theory, one could disentangle them by subsampling the longer time-series to have the same length as the shorter time-series. Clements and Ozgul (2016) studied how the forcing rate affects detectability of early warning signals in simple one-dimensional models. They found that, fixing the length of the time-series, a higher forcing rate led to stronger signals in most of the indicators they considered. When keeping the time-series of the smaller forcing rate at their original, longer length, indicators performed better due to the increased sample size compared to higher forcing rates, which is what we have observed as well. We assumed a linear increase in the bifurcation parameter, and while we expect that our results are robust to different change processes, future work may wish to assess this in more detail.

A principal limitation of our results is that the underlying Generalized Lotka-Volterra model is not an adequate description of the mood system. Therefore, it is not possible to translate our findings as speaking directly to the performance of early warning signals in anticipating critical transitions in the mood system. Instead, our investigation may be interpreted in two ways.

First, our simulation study illustrates how one can investigate the performance of early warning signals to anticipate critical transitions when a model is available. If the model one simulates from is a sensible representation of a psychological system, using ROC curves to quantify the performance of various early warning indicators under different noise intensities, sampling frequencies, baseline periods, and transition periods could then directly inform empirical applications. A ROC curve allows one to see how different decision thresholds balance the true positive and false positive rates. In our investigation, we have focused on the widely used 2σ -threshold, which generally provides a good balance between true positives and false positives. When studying a real psychological system — and if one has a particular intervention in mind — one may instead choose a decision threshold that allows more false positives in order to increase the number of true positive. Lastly, while some of our conclusions must be interpreted with caution as being potentially specific to the Lotka-Volterra model — for example, the finding that skewness and spatial variance performed poorly (see Appendix E) — others, such as the fact that performance decreased with increasing noise or with decreasing sampling frequency, likely hold for a large class of real psychological systems.

The notion of critical transitions between “healthy” and “unhealthy” stable states has played an important role in theoretical psychology (e.g., Borsboom, 2017; Cramer et al., 2016), and a second interpretation of our results understands the Lotka-Volterra model in this way: as an ideal dynamical system that shows critical transitions between two stable states (see also Haslbeck & Ryan, 2021; van de Leemput et al., 2014). In contrast to this ideal system, real psychological systems likely show complications that suppress the occurrence of early warning signals, as discussed in Section 3. For example, the exact components of real psychological systems and how to measure them are generally unknown; it is unclear which components would actually show critical slowing down, if at all; and real psychological systems may be subject to correlated noise, which can suppress early warning indicators. In this interpretation, our finding that early warning indicators are very sensitive to noise in such an ideal system puts into question their potential usefulness in practice. While such a strong inference does not strictly follow — the class of real psychological systems is likely large, and we might find cases where early warning signals perform better than in the model we studied here — we do believe that our results invite a critical discussion on how to move the study of early warning signals forward. We turn to such a discussion in the next section.

5 A Way Forward for Early Warning Signals

Applying early warning signals successfully requires addressing a number of theoretical and practical considerations. In this section, guided by Figure 8, we review these considerations and provide recommendations and suggestions for future research. Several of the points we touch on require

dedicated research programs themselves, and our goal here is not to address them in full; instead, we aim to point out the challenges and choices researchers face when applying early warning signals to anticipate critical transitions in practice.

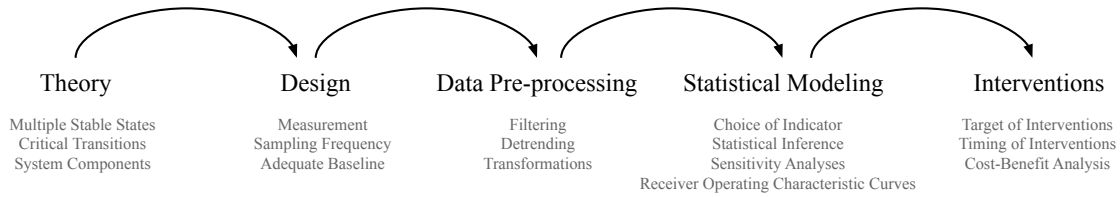


Figure 8: Illustrates the challenges and choices that require careful thought when applying early warning signals in practice.

Theory. In sharp contrast to their reputation of being generic or model-agnostic, we have seen that the occurrence of early warning signals depends on the specifics of the system under study; there are systems that show critical slowing down even though they do not exhibit critical transitions (Kéfi et al., 2013; Wagner & Eisenman, 2015), and there are (a potentially large class of) systems which show critical transitions but no critical slowing down (Hastings & Wysham, 2010). Even if the target system falls into the class of systems that show critical slowing down before critical transitions, early warning signals may only be observable in a small number of variables of the system (Boerlijst et al., 2013; Patterson et al., 2021). While recent work tries to find the system components that most strongly express critical slowing down (Dakos, 2018; Weinans et al., 2019), purely statistical work will not put early warning signals on a solid footing.

Instead, we need to build a basic understanding of the system under study. Does the system actually exhibit abrupt transitions between stable states, or are they smooth? In case of the latter, early warning signals, even if they occur, may not add much predictive utility above other indicators such as the mean. Understanding the types of transitions requires more descriptive research (e.g., Helmich et al., 2020; Hosenfeld et al., 2015; Olthof et al., 2020), an important first step in theory construction (Borsboom et al., 2021). Yet such descriptive research should not only look at a system’s state variables, but relate changes in these to underlying drivers (e.g., Bestelmeyer et al., 2011). In the context of our mood example, this would require measuring the underlying variable “stress” that drives critical transitions in the mood variables, rather than measuring only the latter. Is the relationship between drivers and state variables linear or nonlinear? Early warning signals of critical transitions are intricately linked to a nonlinear relationship — otherwise small changes in a driver would not be able to lead to large changes in the state variables — and studies that establish a nonlinear relationship are indeed much more likely to find evidence of early warning signals (Litzow & Hunsicker, 2016). A necessary condition for a nonlinear relationship that gives rise to multiple stable states and critical transitions is the occurrence of (strong) reinforcing feedback loops (Kéfi et al., 2016; Meadows, 2008). Theoretically mapping the reinforcing feedback loops that may occur in a particular psychological system is therefore an important endeavour. While empirically establishing a nonlinear relationship may turn out to be difficult (e.g., Capon et al., 2015; Dudney & Suding, 2020; Hillebrand et al., 2020; Petraitis, 2013), it would constitute a crucial step forward. Another key question concerns what components constitute the system. Following Borsboom (2017), mental disorders such as major depression are widely viewed not as a latent common cause, but instead arise out of a network of mutually enforcing components; what are these components exactly? Formal models aid in answering these questions by forcing researchers to be explicit in their assumptions about the system (e.g., Robinaugh et al., 2020).

An instructive example of the power of formal modeling for early warning signals comes from epidemiology. O’Regan and Drake (2013) showed that early warning indicators were highly predictive of disease elimination by increased vaccination uptake in the basic Susceptible-Infectious-Susceptible (SIS) and Susceptible-Infectious-Removed (SIR) compartmental models, but not when the disease reemerged as vaccination uptake decreased.⁵ Moreover, while the autocorrelation increased as the disease approached elimination for both models, the variance *decreased* for the SIR

⁵SIS and SIR models segment the population into susceptible (S), infectious (I), and — in the SIR model, which assumes lifelong immunity — due to recovery or death removed (R) groups through which an infectious disease can spread (see e.g., Keeling & Rohani, 2011).

model, which models diseases with life-long immunity. Along similar lines, O'Regan et al. (2016) studied the elimination of vector-borne diseases such as malaria through gradually deployed control measures. They showed that in theory, critical slowing down is expected to occur. In particular, they found that the autocorrelation and variance increased when eliminating malaria through a reduction in the biting rate of the mosquitoes or their population size. However, the autocorrelation did not anticipate elimination through a reduction of the human infectious period or a reduction of the per-capita mosquito mortality rate, while a *decrease* in variance anticipated elimination in both cases. In a similarly impressive modeling feat, O'Dea et al. (2018) showed that, in a SIR model, the autocorrelation of the number of infected provided a better estimate of the distance to the epidemic threshold than the autocorrelation of the number of susceptibles. In recent work combining empirical analysis and theoretical modelling, Dablander et al. (2021) found that early warning indicators tended to *decrease* rather than *increase* prior to the second COVID-19 wave. This is because the underlying driver did not change slowly compared to the characteristic time-scale of the system (a key assumption of critical slowing down), obscuring early warning indicators. All these theoretical insights guide what one may expect in practical applications; they can help us understand why we see, or not see, certain patterns in data. Another example comes from epilepsy research, where models suggested that the change in brain state from normal to seizure can be described as a bifurcation; empirical evidence of critical slowing down broadly supports the use of early warning signals as a means to anticipate seizures (Maturana et al., 2020).

There is little doubt that psychology in general and psychopathology in particular are far removed from the deep understanding epidemiologists have about the mechanics of infectious disease transmission. But this is no reason to despair, with recent modeling efforts in psychopathology paving the way for a future rich in formal modeling (e.g., Bayani et al., 2017; Burger et al., 2020; Cramer et al., 2016; Duncan et al., 2019; Kossakowski et al., 2019; Robinaugh et al., 2019; Schiepek et al., 2017). An approach that lies between purely statistical and sophisticated theoretical modeling work is to tailor relatively simple dynamical system models to the specific phenomenon under study. If one wants to assess whether the target system exhibits bistability, critical transitions, and critical slowing down, then a simple model which allows for these properties is the cusp-catastrophe model (e.g., Zahler & Sussmann, 1977; Zeeman, 1976), which has been fruitfully applied, for example, in the study of attitudes (e.g., van der Maas et al., 2020; van der Maas et al., 2003). In addition to the relative ease with which it can be tested on data (Grasman et al., 2009), the cusp-catastrophe model requires the specification of two underlying drivers, thus moving away from simple statistical models towards incorporating and subsequently testing theoretical assumptions.

Design. Once a basic theoretical understanding of the system has been established, a key question is how to design studies to investigate potential transitions of the system empirically. The most promising empirical application of early warning signals is in online monitoring settings, which come with unique challenges. To shed light on these challenges, we simulated data from a four-dimensional Generalized Lotka-Volterra model, mimicking what using early warning indicators as tools for personalized interventions requires: a way to assess, in an online fashion, whether a person is likely to rapidly transition into an undesired stable state. By sequentially simulating from the model, and at each time point assessing whether the system is likely to transition, we have mirrored this situation in silico. We have found that univariate early warning indicators such as autocorrelation, variance as well as multivariate indicators such as the cross-correlation and the dominant eigenvalue of the covariance matrix perform exceptionally well under low-noise settings. Yet by increasing the noise — to settings that may still be smaller than what we observe in real psychological systems — their predictive performance diminished rapidly. For practical applications, this implies that researchers should do their best to reduce noise. Since measurement error can add to noise, psychometric techniques such as latent variable modeling that reduce measurement error may be a fruitful avenue (Bollen, 1989), as are models for transitions that explicitly take measurement error into account (see e.g., Hefley et al., 2013).

Our simulation results also illustrate other important factors for the design of empirical studies of early warning signals. While we found that as the sampling frequency decreased, the performance of all early warning indicators decreased, this decrease was nonlinear: performance barely dropped when going from sampling ten times a day to five times a day, but then dropped substantially when sampling only once a day.⁶ Collecting data too frequently is an unnecessary burden on the

⁶We want to stress again that it is not a reduction in sampling frequency *by itself* that reduces performance; instead, fixing parameters that are generally not under the control of the researcher (such as noise intensity and the

participant and is likely to reduce compliance, and so an understanding of the time scale of the system is important for the design of empirical studies. For example, if changes happen on the level of weeks, it may be unnecessary to collect data on a daily level. Yet if changes happen on the level of days, collecting data only weekly will be inadequate. The required sampling frequency is crucially dependent on the time it takes the system to reach the tipping point. As we have seen in our simulation results, as the transition period decreased — that is, as resilience was eroded more quickly — detecting early warning signals became more difficult. Understanding the way and speed with which resilience is eroded in, for example, different psychiatric disorders therefore has implications for their empirical study. Similarly, online monitoring requires one to test whether the current early warning indicator has increased compared to a baseline where no or little change in the relevant underlying parameters occurs. As illustrated in our simulation results, the extent of the baseline influences the performance of early warning signals; understanding when the resilience of the system is likely being eroded and when it is not is therefore important in practice.

One core concern that we did not illustrate with our simulation study is measurement: how can we best measure the variables that constitute the system components in practice? In our simulation study, we have assumed a one-to-one mapping between psychological processes (e.g., the mood state “cheerful”) and the observed variables. This is likely not the case in practice; for example, one might observe sudden transitions between positive and negative states when they are assessed momentarily, but smooth transitions when they are assessed retrospectively due to some kind of averaging process that is happening in retrospective assessment. This may have implications for the occurrence of early warning signals, and further stresses the interplay between theorizing and the design of empirical studies (see also Haslbeck et al., [in press](#)).

Data pre-processing. In our simulation setup, the only system-external factor influencing the generation of the data was uncorrelated Gaussian noise, which resulted in data that is in some sense “ideal”. real-world data, on the other hand, is much messier. There may be nonstationarities in the mean or seasonal trends that impose a strong correlational structure on the data, and these can lead to spurious alarms or missed transitions. Data pre-processing steps, such as detrending to remove nonstationarities in the mean, and filtering to remove potential seasonalities, are important aspects in the study of early warning signals (Dakos et al., 2012a; Jäger & Füllsack, 2019; Lenton et al., 2012). These come in many variants, however, and therefore either afford a theoretically motivated choice or — more likely — a thorough sensitivity analyses. The majority of studies investigating early warning signals in fields outside of psychology provide such a sensitivity analysis by varying, for example, the bandwidth of a Gaussian filter and reporting for which settings a transition is indicated. This is currently not being considered in the psychology literature, and we suggest that researchers routinely employ such sensitivity analyses to avoid drawing potentially spurious conclusions. In online monitoring settings, such sensitivity analyses would result in a distribution over early warning indicators, requiring the definition of thresholds beyond which one would signal a transition in practice.

Statistical modeling. Once data has been collected and pre-processed, the question turns to statistical modeling. First, what early warning indicators should we employ? The list of suggested indicators is large, ranging from simple univariate indicators such as autocorrelation, variance, kurtosis, and skewness (Carpenter & Brock, 2006; Guttal & Jayaprakash, 2008; Scheffer et al., 2009) to multivariate indicators such as spatial variance and cross-correlation (e.g., Dakos et al., 2010; Kéfi et al., 2014; Weinans et al., 2021). These can be implemented in a nonparametric way using rolling windows, or in the form of parametric time-varying models (Dakos et al., 2012a). In our simulation study, we have focused on the nonparametric approach due to its ubiquity and ease of use — virtually any summary statistic can be computed by means of rolling windows, while implementing and estimating them in parametric models can be challenging. The size of the rolling window is an important parameter: small window sizes lead to noisier estimates, while large window sizes lead to a reduced sensitivity to detect changes in the indicator. Understanding the extent to which a difference in rolling window sizes influences early warning signals is therefore prudent. While such sensitivity analyses are widespread in the study of early warning signals in fields outside of psychology, they are absent in psychology itself. Thus, to avoid drawing potentially spurious conclusions, we not only recommend doing sensitivity analyses in the pre-processing of

transition period), a reduction in sampling frequency results in fewer observations, a more detrimental impact of noise, a less fine-grained picture of the change process, thereby reducing performance.

data, but also in the statistical analysis when relying on rolling windows. Similarly, because single indicators are more prone to false alarms (e.g., Ditlevsen & Johnsen, 2010; Wagner & Eisenman, 2015), we recommend combining several indicators using the methodology we outlined in our simulation study.

Other statistical modeling approaches incorporate assumptions into models and then select between the models. For example, Boettiger and Hastings (2012b) suggested comparing two models, one of which assumes that the system undergoes a saddle-node bifurcation and another one that does not, while Bury et al. (2021) used deep learning to train a classifier to distinguish between three types of bifurcations and a system that undergoes no transition, finding that their method strongly outperformed traditional univariate indicators. Lade and Gross (2012) outlined an approach which builds assumptions about the dynamics of the system into the model, finding that incorporating such information can lead to improved performance. This approach moves away from the generic aspect of indicators towards incorporating features that are specific to the particular system under study. The call to resist the lure of elusive “generic” early warning signals is not new in ecology (Boettiger & Hastings, 2013), but has become stronger as their limitations have become more apparent (Clements & Ozgul, 2018). We believe that resisting this lure is important also in a psychological context, and that there are exciting avenues for future research that incorporates system-specific information in order to better anticipate transitions in psychological systems.

Lastly, if a model of a target system — even if simplified — is available, researchers should assess the performance of their statistical method in simulation by using, as we have illustrated, ROC curves to understand the true positive and false positive rate of the method for various decision thresholds (see also for example Clements et al., 2019).

Interventions. Ultimately, the potential of early warning signals lies not only in anticipating critical transitions, but also in allowing us to intervene in time to avoid such transitions. Several researchers have suggested that early warning signals could be very useful for personalized early intervention in the context of, for example, psychotherapy and psychiatric disorders (e.g., Hayes et al., 2015; Hofmann et al., 2016; Nelson et al., 2017; Olthof et al., 2019; Schiepek et al., 2011; Wichers et al., 2019; Wichers et al., 2020). A key issue is how exactly we would go about intervening in the system, and what the respective costs and benefits are — these are needed for a decision-theoretic analysis that goes beyond mere statistical inference.

Another key issue is whether we can actually intervene in time to avert a critical transition. While we did not study interventions in our simulation study directly, our results still offer some insight. In particular, whether we can intervene in time to avert a transition depends on how far in advance early warning indicators signal a transition. We have seen that early warning indicators can signal a transition far in advance of the actual transition if noise is small. Under more realistic, larger noise settings early warning signals occur closer to the actual transition. Suppose one wishes to use early warning signals to anticipate the onset of depression, and suppose one observes a significant increase in indicators; how far is the transition away? How much time do we have to intervene and avoid a transition? Since this depends strongly on the quality of the data and on the speed with which the system approaches the tipping point, relative changes in indicators cannot provide an answer. Instead, absolute levels of early warning indicators may be used to estimate the distance to the tipping point (O’Dea et al., 2018; Wissel, 1984). However, these absolute levels will depend on the system under study as well as on specifics of the data. The extent of the autocorrelation, for example, reflects the ratio of the time scale of the system to the sampling frequency. To illustrate, suppose the system changes slowly, say on the level of weeks. If one samples it many times a day, then the measurements will exhibit a high autocorrelation. While the autocorrelation technically approaches one at the tipping point (more specifically, at the zero-eigenvalue bifurcation), this is the case only for unidimensional systems. For coupled multidimensional systems — and thus, all systems of real-world interest — the level of the autocorrelation of different system components at the tipping point is system-dependent, and so using absolute levels to assess the distance to the tipping point requires sophisticated understanding (e.g., O’Dea et al., 2018). Biggs et al. (2009) found that relative changes in autocorrelation and variance can occur too late to avert a transition in a fisheries model, but that absolute levels in system-specific indicators would enable one to act in time. Thus, similar to what we have seen in the context of anticipating transitions, using early warning signals as a way to intervene in time to avert a transition requires a good understanding of the target system.

If the goal is to anticipate a transition into an unhealthy state, then early warning signals

alone may not be the best way to advance such a goal. Take the onset of major depression as a prediction problem, for example. Are moods momentarily assessed (e.g., Wichers et al., 2020) really the most predictive features? One might go beyond early warning indicators by using information from sensor data, including physiological variables and even geo-location data (e.g., Saeb et al., 2016). This approach might draw from the growing literature on machine learning in psychiatry and psychopathology (e.g., Bzdok & Meyer-Lindenberg, 2018; Dwyer et al., 2018; Mohr et al., 2017). For example, Webb et al. (2020) tried to predict post-treatment depression scores in a clinical sample that underwent treatment in a psychiatric hospital using a range of pre-treatment variables. While they did not anticipate the onset of depressive episodes in real-time, one may well translate their approach to further such a goal. Along these lines, Jacobson and Chung (2020) used sensor data — including data on location, weather, and heart rate — collected using smartphones in the past 24 hours to predict the next hour of depression symptom severity in a small sample reporting clinical levels of depression. These are important first studies indicating that a predictive approach that goes beyond generic early warning signals may be feasible; if the goal is simply to predict the onset of mental disorders such as depression, this strikes us as a viable approach well worth exploring. Such an approach does not aim to develop formal models of mental disorders, however, and this means a lack of mechanistic understanding. It is also less likely to inform promising targets of interventions. Formal models, on the other hand, have the potential to inform interventions and may also result in more powerful, system-specific early warning indicators of transitions. Harnessing the full potential of early warning signals, then, is deeply intertwined with the goal of advancing our understanding of mental disorders by building formal models.

6 Conclusion

Dynamical systems theory provides a unifying framework for studying how systems as disparate as the climate and the behaviour of humans change over time. Some systems can exhibit multiple stable states and critical transitions between them due to internally changing dynamics. Early warning signals based on critical slowing down that may anticipate such transitions have been widely discussed in ecology, epidemiology, and climate science in the last two decades, finding shelter also in psychology and psychiatry. With this paper, we hope to contribute to a better understanding of their limitations; to an appreciation of the challenges associated with their use in practice; and to possible avenues that may put the study of early warning signals in psychology and psychiatry on a solid footing.

Author Contributions. F. Dablander suggested the research topic and all authors have been involved in the design of the work. F. Dablander implemented the simulation study, analyzed the data, and wrote the paper. A. Pichler added the mathematical derivations. All authors have been involved in the interpretation of the results, proof-read, and approved the submitted version of the paper. They also declare that there were no conflicts of interest.

Acknowledgements. We thank the organizers of the Complex Systems Summer School hosted at the Institute of American Indian Arts and the Santa Fe Institute, where this research originated. We thank Jonas Haslbeck and Oisín Ryan for valuable comments on various earlier versions of this manuscript. We also thank Tessa Blanken, Denny Borsboom, Julian Burger, Max Hinne, Han van der Maas, and Don Robinaugh for valuable comments on an earlier version of this manuscript.

Materials. All figures are licensed under a CC BY 4.0 license and are available from <https://github.com/fdabl/Early-Warning>, which also includes the code to reproduce the simulation, analyses, and figures.

References

Ashwin, P., Wieczorek, S., Vitolo, R., & Cox, P. (2012). Tipping points in open systems: Bifurcation, noise-induced and rate-dependent examples in the climate system. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 370(1962), 1166–1184.

- Bayani, A., Hadaeghi, F., Jafari, S., & Murray, G. (2017). Critical slowing down as an early warning of transitions in episodes of bipolar disorder: A simulation study based on a computational model of circadian activity rhythms. *Chronobiology International*, *34*(2), 235–245.
- Bestelmeyer, B. T., Ellison, A. M., Fraser, W. R., Gorman, K. B., Holbrook, S. J., Laney, C. M., Ohman, M. D., Peters, D. P., Pillsbury, F. C., Rassweiler, A., Et al. (2011). Analysis of abrupt transitions in ecological systems. *Ecosphere*, *2*(12), 1–26.
- Biggs, R., Carpenter, S. R., & Brock, W. A. (2009). Turning back from the brink: Detecting an impending regime shift in time to avert it. *Proceedings of the National Academy of Sciences*, *106*(3), 826–831.
- Boerlijst, M. C., Oudman, T., & de Roos, A. M. (2013). Catastrophic collapse can occur without early warning: Examples of silent catastrophes in structured ecological models. *PLoS One*, *8*(4), e62033.
- Boettiger, C. (2018). From noise to knowledge: How randomness generates novel phenomena and reveals information. *Ecology Letters*, *21*(8), 1255–1267.
- Boettiger, C., & Hastings, A. (2012a). Early warning signals and the prosecutor’s fallacy. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1748), 4734–4739.
- Boettiger, C., & Hastings, A. (2012b). Quantifying limits to detection of early warning for critical transitions. *Journal of the Royal Society Interface*, *9*(75), 2527–2539.
- Boettiger, C., & Hastings, A. (2013). From patterns to predictions. *Nature*, *493*(7431), 157–158.
- Boettiger, C., Ross, N., & Hastings, A. (2013). Early warning signals: The charted and uncharted territories. *Theoretical Ecology*, *6*(3), 255–264.
- Bollen, K. A. (1989). *Structural Equations with Latent Variables*. Wiley.
- Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry*, *16*(1), 5–13.
- Borsboom, D., van der Maas, H., Dalege, J., Kievit, R., & Haig, B. (2021). Theory construction methodology: A practical framework for building theories in psychology. *Perspectives on Psychological Science*, *16*, 756–766.
- Boulton, C. A., Good, P., & Lenton, T. M. (2013). Early warning signals of simulated Amazon rainforest dieback. *Theoretical Ecology*, *6*(3), 373–384.
- Brett, T. S., O’Dea, E. B., Marty, É., Miller, P. B., Park, A. W., Drake, J. M., & Rohani, P. (2018). Anticipating epidemic transitions with imperfect data. *PLoS Computational Biology*, *14*(6), e1006204.
- Burger, J., van der Veen, D. C., Robinaugh, D. J., Quax, R., Riese, H., Schoevers, R. A., & Epskamp, S. (2020). Bridging the gap between complexity science and clinical practice by formalizing idiographic theories: A computational model of functional analysis. *BMC Medicine*, *18*, 1–18.
- Bury, T. M., Bauch, C. T., & Anand, M. (2020). Detecting and distinguishing tipping points using spectral early warning signals. *Journal of the Royal Society Interface*, *17*(170), 20200482.
- Bury, T., Sujith, R., Pavithran, I., Scheffer, M., Lenton, T., Anand, M., & Bauch, C. (2021). Deep learning for early warning signals of tipping points. *Proceedings of the National Academy of Sciences*, *118*(39), e2106140118.
- Bzdok, D., & Meyer-Lindenberg, A. (2018). Machine learning for precision psychiatry: Opportunities and challenges. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *3*(3), 223–230.
- Capon, S. J., Lynch, A. J. J., Bond, N., Chessman, B. C., Davis, J., Davidson, N., Finlayson, M., Gell, P. A., Hohnberg, D., Humphrey, C., Et al. (2015). Regime shifts, thresholds and multiple stable states in freshwater ecosystems; a critical appraisal of the evidence. *Science of the Total Environment*, *534*, 122–130.
- Carpenter, S. R., & Brock, W. A. (2006). Rising variance: A leading indicator of ecological transition. *Ecology Letters*, *9*(3), 311–318.
- Chen, S., O’Dea, E. B., Drake, J. M., & Epureanu, B. I. (2019). Eigenvalues of the covariance matrix as early warning signals for critical transitions in ecological systems. *Scientific Reports*, *9*(1), 2572.
- Chevance, A., Ravaud, P., Tomlinson, A., Le Berre, C., Teufer, B., Touboul, S., Fried, E. I., Gartlehner, G., Cipriani, A., & Tran, V. T. (2020). Identifying outcomes for depression that matter to patients, informal caregivers, and health-care professionals: Qualitative content analysis of a large international online survey. *The Lancet Psychiatry*, *7*(8), 692–702.

- Chisholm, R. A., & Filotas, E. (2009). Critical slowing down as an indicator of transitions in two-species models. *Journal of Theoretical Biology*, *257*(1), 142–149.
- 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.
- Clements, C. F., McCarthy, M. A., & Blanchard, J. L. (2019). Early warning signals of recovery in complex systems. *Nature Communications*, *10*(1), 1–9.
- Clements, C. F., & Ozgul, A. (2016). Rate of forcing and the forecastability of critical transitions. *Ecology and Evolution*, *6*(21), 7787–7793.
- Clements, C. F., & Ozgul, A. (2018). Indicators of transitions in biological systems. *Ecology Letters*, *21*(6), 905–919.
- Cramer, A. O., van Borkulo, C. D., Giltay, E. J., van der Maas, H. L., Kendler, K. S., Scheffer, M., & Borsboom, D. (2016). Major depression as a complex dynamic system. *PLoS One*, *11*(12), e0167490.
- Curtiss, J., Fulford, D., Hofmann, S. G., & Gershon, A. (2019). Network dynamics of positive and negative affect in bipolar disorder. *Journal of Affective Disorders*, *249*, 270–277.
- Dablander, F. (2020). A gentle introduction to dynamical systems theory [Accessed: 07/10/2021]. <https://fabindablander.com/r/Dynamical-Systems.html>
- Dablander, F., Heesterbeek, H., Borsboom, D., & Drake, J. M. (2021). Overlapping Time Scales Obscure Early Warning Signals of the Second COVID-19 Wave. <https://doi.org/10.1101/2021.07.27.21261226>
- Dai, L., Korolev, K. S., & Gore, J. (2013). Slower recovery in space before collapse of connected populations. *Nature*, *496*(7445), 355–358.
- Dai, L., Korolev, K. S., & Gore, J. (2015). Relation between stability and resilience determines the performance of early warning signals under different environmental drivers. *Proceedings of the National Academy of Sciences*, *112*(32), 10056–10061.
- Dai, L., Vorselen, D., Korolev, K. S., & Gore, J. (2012). Generic indicators for loss of resilience before a tipping point leading to population collapse. *Science*, *336*(6085), 1175–1177.
- Dakos, V. (2018). Identifying best-indicator species for abrupt transitions in multispecies communities. *Ecological Indicators*, *94*, 494–502.
- 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. (2012a). Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data. *PLoS One*, *7*(7), e41010.
- Dakos, V., Scheffer, M., van Nes, E. H., Brovkin, V., Petoukhov, V., & Held, H. (2008). Slowing down as an early warning signal for abrupt climate change. *Proceedings of the National Academy of Sciences*, *105*(38), 14308–14312.
- Dakos, V., van Nes, E. H., D’Odorico, P., & Scheffer, M. (2012b). Robustness of variance and autocorrelation as indicators of critical slowing down. *Ecology*, *93*(2), 264–271.
- Dakos, V., van Nes, E. H., Donangelo, R., Fort, H., & Scheffer, M. (2010). Spatial correlation as leading indicator of catastrophic shifts. *Theoretical Ecology*, *3*(3), 163–174.
- Ditlevsen, P. D., & Johnsen, S. J. (2010). Tipping points: Early warning and wishful thinking. *Geophysical Research Letters*, *37*(19).
- Drake, J. M., & Griffen, B. D. (2010). Early warning signals of extinction in deteriorating environments. *Nature*, *467*(7314), 456–459.
- Dudney, J., & Suding, K. N. (2020). The elusive search for tipping points. *Nature Ecology & Evolution*, *4*, 1449–1450.
- Duncan, J. P., Aubele-Futch, T., & McGrath, M. (2019). A fast-slow dynamical system model of addiction: Predicting relapse frequency. *SIAM Journal on Applied Dynamical Systems*, *18*(2), 881–903.
- Dwyer, D. B., Falkai, P., & Koutsouleris, N. (2018). Machine learning approaches for clinical psychology and psychiatry. *Annual Review of Clinical Psychology*, *14*, 91–118.
- Gardiner, C. (2004). *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences* (Vol. 3). Springer.
- Gijzel, S. M., Whitson, H. E., van de Leemput, I. A., Scheffer, M., van Asselt, D., Rector, J. L., Olde Rikkert, M. G., & Melis, R. J. (2019). Resilience in clinical care: Getting a grip on the recovery potential of older adults. *Journal of the American Geriatrics Society*, *67*(12), 2650–2657.

- Grasman, R., van der Maas, H. L., & Wagenmakers, E.-J. (2009). Fitting the Cusp Catastrophe in R: A cusp Package Primer. *Journal of Statistical Software*, *32*(1), 1–27.
- Guttal, V., & Jayaprakash, C. (2008). Changing skewness: An early warning signal of regime shifts in ecosystems. *Ecology Letters*, *11*(5), 450–460.
- Guttal, V., & Jayaprakash, C. (2009). Spatial variance and spatial skewness: Leading indicators of regime shifts in spatial ecological systems. *Theoretical Ecology*, *2*(1), 3–12.
- Hamilton, J. D. (1994). *Time Series Analysis* (Vol. 2). Princeton University Press.
- Harris, M. J., Hay, S. I., & Drake, J. M. (2020). Early warning signals of malaria resurgence in Kericho, Kenya. *Biology Letters*, *16*(3), 20190713.
- Haslbeck, J. M., & Ryan, O. (2021). Recovering within-person dynamics from psychological time series. *Multivariate Behavioral Research*, 1–32.
- Haslbeck, J., Ryan, O., Robinaugh, D., Waldorp, L., & Borsboom, D. (in press). Modeling Psychopathology: From Data Models to Formal Theories. *Psychological Methods*.
- Hastings, A., & Wysham, D. B. (2010). Regime shifts in ecological systems can occur with no warning. *Ecology Letters*, *13*(4), 464–472.
- Hayes, A. M., & Andrews, L. A. (2020). A complex systems approach to the study of change in psychotherapy. *BMC Medicine*, *18*(1), 1–13.
- Hayes, A. M., Laurenceau, J.-P., Feldman, G., Strauss, J. L., & Cardaciotto, L. (2007). Change is not always linear: The study of nonlinear and discontinuous patterns of change in psychotherapy. *Clinical Psychology Review*, *27*(6), 715–723.
- Hayes, A. M., Yasinski, C., Barnes, J. B., & Bockting, C. L. (2015). Network destabilization and transition in depression: New methods for studying the dynamics of therapeutic change. *Clinical Psychology Review*, *41*, 27–39.
- Hayes, S. C., Hofmann, S. G., Stanton, C. E., Carpenter, J. K., Sanford, B. T., Curtiss, J. E., & Ciarrochi, J. (2019). The role of the individual in the coming era of process-based therapy. *Behaviour Research and Therapy*, *117*, 40–53.
- Hefley, T. J., Tyre, A. J., & Blankenship, E. E. (2013). Statistical indicators and state–space population models predict extinction in a population of bobwhite quail. *Theoretical Ecology*, *6*(3), 319–331.
- Helmich, M. A., Wichers, M., Olthof, M., Strunk, G., Aas, B., Aichhorn, W., Schiepek, G., & Snippe, E. (2020). Sudden gains in day-to-day change: Revealing nonlinear patterns of individual improvement in depression. *Journal of Consulting and Clinical Psychology*, *88*(2), 119–127.
- Hillebrand, H., Donohue, I., Harpole, W. S., Hodapp, D., Kucera, M., Lewandowska, A. M., Merder, J., Montoya, J. M., & Freund, J. A. (2020). Thresholds for ecological responses to global change do not emerge from empirical data. *Nature Ecology & Evolution*, *4*, 1502–1509.
- Hofmann, S. G., Curtiss, J., & McNally, R. J. (2016). A complex network perspective on clinical science. *Perspectives on Psychological Science*, *11*(5), 597–605.
- Holling, C. S. (1973). Resilience and stability of ecological systems. *Annual Review of Ecology, Evolution, and Systematics*, *4*(1), 1–23.
- Hosenfeld, B., Bos, E. H., Wardenaar, K. J., Conradi, H. J., van der Maas, H. L., Visser, I., & de Jonge, P. (2015). Major depressive disorder as a nonlinear dynamic system: Bimodality in the frequency distribution of depressive symptoms over time. *BMC Psychiatry*, *15*(1), 222.
- Jacobson, N. C., & Chung, Y. J. (2020). Passive sensing of prediction of moment-to-moment depressed mood among undergraduates with clinical levels of depression sample using smartphones. *Sensors*, *20*(12), 3572.
- Jäger, G., & Füllsack, M. (2019). Systematically false positives in early warning signal analysis. *PLoS One*, *14*(2), e0211072.
- Kalisch, R., Cramer, A. O., Binder, H., Fritz, J., Leertouwer, I., Lunansky, G., Meyer, B., Timmer, J., Veer, I. M., & Van Harmelen, A.-L. (2019). Deconstructing and reconstructing resilience: A dynamic network approach. *Perspectives on Psychological Science*, *14*(5), 765–777.
- Keeling, M. J., & Rohani, P. (2011). *Modeling Infectious Diseases in Humans and Animals*. Princeton University Press.
- Kéfi, S., Dakos, V., Scheffer, M., van Nes, E. H., & Rietkerk, M. (2013). Early warning signals also precede non-catastrophic transitions. *Oikos*, *122*(5), 641–648.

- Kéfi, S., Guttal, V., Brock, W. A., Carpenter, S. R., Ellison, A. M., Livina, V. N., Seekell, D. A., Scheffer, M., van Nes, E. H., & Dakos, V. (2014). Early warning signals of ecological transitions: Methods for spatial patterns. *PLoS One*, *9*(3), e92097.
- Kéfi, S., Holmgren, M., & Scheffer, M. (2016). When can positive interactions cause alternative stable states in ecosystems? *Functional Ecology*, *30*(1), 88–97.
- Killick, R., & Eckley, I. (2014). changepoint: An R package for changepoint analysis. *Journal of Statistical Software*, *58*(3), 1–19.
- Kossakowski, J. J., Gordijn, M., Riese, H., & Waldorp, L. J. (2019). Applying a dynamical systems model and network theory to major depressive disorder. *Frontiers in Psychology*, *10*, 1762.
- Kramer, A. M., Dennis, B., Liebhold, A. M., & Drake, J. M. (2009). The evidence for Allee effects. *Population Ecology*, *51*(3), 341–354.
- Kuranova, A., Booij, S. H., Menne-Lothmann, C., Decoster, J., van Winkel, R., Delespaul, P., De Hert, M., Derom, C., Thiery, E., Rutten, B. P., Et al. (2020). Measuring resilience prospectively as the speed of affect recovery in daily life: A complex systems perspective on mental health. *BMC Medicine*, *18*(1), 1–11.
- Lade, S. J., & Gross, T. (2012). Early warning signals for critical transitions: A generalized modeling approach. *PLoS Computational Biology*, *8*(2), e1002360.
- Lenton, T. M. (2011). Early warning of climate tipping points. *Nature Climate Change*, *1*(4), 201–209.
- Lenton, T. M., Livina, V., Dakos, V., van Nes, E., & Scheffer, M. (2012). Early warning of climate tipping points from critical slowing down: Comparing methods to improve robustness. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *370*(1962), 1185–1204.
- Litzow, M. A., & Hunsicker, M. E. (2016). Early warning signals, nonlinearity, and signs of hysteresis in real ecosystems. *Ecosphere*, *7*(12), e01614.
- Lohmann, J., & Ditlevsen, P. D. (2021). Risk of tipping the overturning circulation due to increasing rates of ice melt. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(9), e2017989118.
- Luke, C., & Cox, P. (2011). Soil carbon and climate change: from the Jenkinson effect to the compost-bomb instability. *European Journal of Soil Science*, *62*(1), 5–12.
- Maturana, M. I., Meisel, C., Dell, K., Karoly, P. J., D’Souza, W., Grayden, D. B., Burkitt, A. N., Jiruska, P., Kudlacek, J., Hlinka, J., Et al. (2020). Critical slowing down as a biomarker for seizure susceptibility. *Nature Communications*, *11*(1), 1–12.
- May, R. M. (1977). Thresholds and breakpoints in ecosystems with a multiplicity of stable states. *Nature*, *269*(5628), 471–477.
- Meadows, D. H. (2008). *Thinking in Systems: A primer*. Chelsea Green Publishing.
- Milkoreit, M., Hodbod, J., Baggio, J., Benessaiah, K., Calderón-Contreras, R., Donges, J. F., Mathias, J.-D., Rocha, J. C., Schoon, M., & Werners, S. E. (2018). Defining tipping points for social-ecological systems scholarship—an interdisciplinary literature review. *Environmental Research Letters*, *13*(3), 033005.
- Mohr, D. C., Zhang, M., & Schueller, S. M. (2017). Personal sensing: Understanding mental health using ubiquitous sensors and machine learning. *Annual Review of Clinical Psychology*, *13*, 23–47.
- Nelson, B., McGorry, P. D., Wichers, M., Wigman, J. T., & Hartmann, J. A. (2017). Moving from static to dynamic models of the onset of mental disorder: A review. *JAMA Psychiatry*, *74*(5), 528–534.
- O’Dea, E. B., Park, A. W., & Drake, J. M. (2018). Estimating the distance to an epidemic threshold. *Journal of the Royal Society Interface*, *15*(143), 20180034.
- Olthof, M., Hasselman, F., & Lichtwarck-Aschoff, A. (2020). Complexity in Psychological Self-Ratings: Implications for research and practice. *BMC Medicine*, *18*(317).
- Olthof, M., Hasselman, F., Strunk, G., van Rooij, M., Aas, B., Helmich, M. A., Schiepek, G., & Lichtwarck-Aschoff, A. (2019). Critical fluctuations as an early-warning signal for sudden gains and losses in patients receiving psychotherapy for mood disorders. *Clinical Psychological Science*, *8*(1), 25–35.
- O’Regan, S. M., & Burton, D. L. (2018). How Stochasticity Influences Leading Indicators of Critical Transitions. *Bulletin of Mathematical Biology*, *80*(6), 1630–1654.
- O’Regan, S. M., & Drake, J. M. (2013). Theory of early warning signals of disease emergence and leading indicators of elimination. *Theoretical Ecology*, *6*(3), 333–357.

- O'Regan, S. M., Lillie, J. W., & Drake, J. M. (2016). Leading indicators of mosquito-borne disease elimination. *Theoretical Ecology*, *9*(3), 269–286.
- Patterson, A. C., Strang, A. G., & Abbott, K. C. (2021). When and where we can expect to see early warning signals in multispecies systems approaching tipping points: Insights from theory. *The American Naturalist*, *198*(1), E12–E26.
- Peretti, C., & Munch, S. (2012). Regime shift indicators fail under noise levels commonly observed in ecological systems. *Ecological Applications*, *22*(6), 1772–1779.
- Perryman, C. G. (2015). How fast is too fast? Rate-induced bifurcations in multiple time-scale systems (PhD Thesis).
- Petraitis, P. (2013). *Multiple Stable States in Natural Ecosystems*. OUP Oxford.
- Qin, S., & Tang, C. (2018). Early-warning signals of critical transition: Effect of extrinsic noise. *Physical Review E*, *97*(3), 032406.
- Ritchie, P., & Sieber, J. (2016). Early-warning indicators for rate-induced tipping. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *26*(9), 093116.
- Robinaugh, D., Haslbeck, J., Ryan, O., Fried, E. I., & Waldorp, L. (2020). Invisible hands and fine calipers: A call to use formal theory as a toolkit for theory construction.
- Robinaugh, D., Haslbeck, J., Waldorp, L., Kossakowski, J., Fried, E. I., Millner, A., McNally, R. J., van Nes, E. H., Scheffer, M., Kendler, K. S., & Borsboom, D. (2019). Advancing the Network Theory of Mental Disorders: A Computational Model of Panic Disorder. <https://doi.org/10.31234/osf.io/km37w>
- Robinson, A., Calov, R., & Ganopolski, A. (2012). Multistability and critical thresholds of the Greenland ice sheet. *Nature Climate Change*, *2*(6), 429–432.
- Saeb, S., Lattie, E. G., Schueller, S. M., Kording, K. P., & Mohr, D. C. (2016). The relationship between mobile phone location sensor data and depressive symptom severity. *PeerJ*, *4*, e2537.
- Scheffer, M. (2009). *Critical Transitions in Nature and Society*. Princeton University Press.
- Scheffer, M., Bascompte, J., Brock, W. A., Brovkin, V., Carpenter, S. R., Dakos, V., Held, H., van Nes, E. H., Rietkerk, M., & Sugihara, G. (2009). Early-warning signals for critical transitions. *Nature*, *461*(7260), 53–59.
- Scheffer, M., Bolhuis, J. E., Borsboom, D., Buchman, T. G., Gijzel, S. M., Goulson, D., Kammenga, J. E., Kemp, B., van de Leemput, I. A., Levin, S., Et al. (2018). Quantifying resilience of humans and other animals. *Proceedings of the National Academy of Sciences*, *115*(47), 11883–11890.
- Scheffer, M., Carpenter, S. R., Dakos, V., & van Nes, E. H. (2015). Generic indicators of ecological resilience: Inferring the chance of a critical transition. *Annual Review of Ecology, Evolution, and Systematics*, *46*, 145–167.
- Scheffer, M., Carpenter, S. R., Lenton, T. M., Bascompte, J., Brock, W. A., Dakos, V., Van de Koppel, J., Van de Leemput, I. A., Levin, S. A., van Nes, E. H., Et al. (2012). Anticipating critical transitions. *Science*, *338*(6105), 344–348.
- Scheffer, M., Van Nes, E. H., Holmgren, M., & Hughes, T. (2008). Pulse-driven loss of top-down control: The critical-rate hypothesis. *Ecosystems*, *11*(2), 226–237.
- Schiepek, G. K., Viol, K., Aichhorn, W., Hütt, M.-T., Sungler, K., Pincus, D., & Schöller, H. J. (2017). Psychotherapy is chaotic—(Not only) in a computational world. *Frontiers in Psychology*, *8*, 379.
- Schiepek, G., Fartacek, C., Sturm, J., Kralovec, K., Fartacek, R., & Plöderl, M. (2011). Nonlinear dynamics: Theoretical perspectives and application to suicidology. *Suicide and Life-Threatening Behavior*, *41*(6), 661–675.
- Schreuder, M. J., Hartman, C. A., George, S. V., Menne-Lothmann, C., Decoster, J., van Winkel, R., Delespaul, P., De Hert, M., Derom, C., Thiery, E., Et al. (2020). Early warning signals in psychopathology: What do they tell? *BMC Medicine*, *18*(1), 1–11.
- Siteur, K., Eppinga, M. B., Doelman, A., Siero, E., & Rietkerk, M. (2016). Ecosystems off track: Rate-induced critical transitions in ecological models. *Oikos*, *125*(12), 1689–1699.
- Stephens, P. A., & Sutherland, W. J. (1999). Consequences of the Allee effect for behaviour, ecology and conservation. *Trends in Ecology & Evolution*, *14*(10), 401–405.
- Strogatz, S. H. (2015). *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering (2nd edition)*. Colorado, US: Westview Press.
- Titus, M., & Watson, J. (2020). Critical speeding up as an early warning signal of stochastic regime shifts. *Theoretical Ecology*, 1–9.

- Trull, T. J., Lane, S. P., Koval, P., & Ebner-Priemer, U. W. (2015). Affective dynamics in psychopathology. *Emotion Review*, 7(4), 355–361.
- van de Leemput, I. A., Wichers, M., Cramer, A. O., Borsboom, D., Tuerlinckx, F., Kuppens, P., van Nes, E. H., Viechtbauer, W., Giltay, E. J., Aggen, S. H., Et al. (2014). Critical slowing down as early warning for the onset and termination of depression. *Proceedings of the National Academy of Sciences*, 111(1), 87–92.
- van der Bolt, B., & van Nes, E. H. (2021). Understanding the critical rate of environmental change for ecosystems, cyanobacteria as an example. *PLoS One*, 16(6), e0253003.
- van der Maas, H. L., Dalege, J., & Waldorp, L. (2020). The polarization within and across individuals: the hierarchical Ising opinion model. *Journal of Complex Networks*, 8(2), cnaa010.
- van der Maas, H. L., Kolstein, R., & Van Der Pligt, J. (2003). Sudden transitions in attitudes. *Sociological Methods & Research*, 32(2), 125–152.
- van Nes, E. H., Arani, B. M., Staal, A., van der Bolt, B., Flores, B. M., Bathiany, S., & Scheffer, M. (2016). What do you mean, ‘tipping point’? *Trends in Ecology & Evolution*, 31(12), 902–904.
- van Nes, E. H., & Scheffer, M. (2004). Large species shifts triggered by small forces. *The American Naturalist*, 164(2), 255–266.
- van Nes, E. H., & Scheffer, M. (2007). Slow recovery from perturbations as a generic indicator of a nearby catastrophic shift. *The American Naturalist*, 169(6), 738–747.
- Wagner, T. J., & Eisenman, I. (2015). False alarms: How early warning signals falsely predict abrupt sea ice loss. *Geophysical Research Letters*, 42(23), 10–333.
- Webb, C. A., Cohen, Z. D., Beard, C., Forgeard, M., Peckham, A. D., & Björgvinsson, T. (2020). Personalized prognostic prediction of treatment outcome for depressed patients in a naturalistic psychiatric hospital setting: A comparison of machine learning approaches. *Journal of Consulting and Clinical Psychology*, 25–38.
- Weinans, E., Lever, J. J., Bathiany, S., Quax, R., Bascompte, J., van Nes, E. H., Scheffer, M., & van de Leemput, I. A. (2019). Finding the direction of lowest resilience in multivariate complex systems. *Journal of the Royal Society Interface*, 16(159), 1–9.
- Weinans, E., Quax, R., van Nes, E. H., & van de Leemput, I. A. (2021). Evaluating the performance of multivariate indicators of resilience loss. *Scientific Reports*, 11(1), 1–11.
- Wichers, M., Wigman, J., & Myin-Germeys, I. (2015). Micro-level affect dynamics in psychopathology viewed from complex dynamical system theory. *Emotion Review*, 7(4), 362–367.
- Wichers, M., Groot, P. C., Psychosystems, Group, E., & Group, E. (2016). Critical slowing down as a personalized early warning signal for depression. *Psychotherapy and Psychosomatics*, 85(2), 114–116.
- Wichers, M., Schreuder, M. J., Goekoop, R., & Groen, R. N. (2019). Can we predict the direction of sudden shifts in symptoms? transdiagnostic implications from a complex systems perspective on psychopathology. *Psychological Medicine*, 49(3), 380–387.
- Wichers, M., Smit, A. C., & Snippe, E. (2020). Early Warning Signals Based on Momentary Affect Dynamics can Expose Nearby Transitions in Depression: A Confirmatory Single-Subject Time-Series Study. *Journal for Person-Oriented Research*, 6(1), 1–15.
- Wilkinson, G. M., Carpenter, S. R., Cole, J. J., Pace, M. L., Batt, R. D., Buelo, C. D., & Kurtzweil, J. T. (2018). Early warning signals precede cyanobacterial blooms in multiple whole-lake experiments. *Ecological Monographs*, 88(2), 188–203.
- Wissel, C. (1984). A universal law of the characteristic return time near thresholds. *Oecologia*, 65(1), 101–107.
- Zahler, R. S., & Sussmann, H. J. (1977). Claims and accomplishments of applied catastrophe theory. *Nature*, 269(5631), 759–763.
- Zeeman, E. C. (1976). Catastrophe theory. *Scientific American*, 234(4), 65–83.

A Critical Slowing Down for Unidimensional Systems

Here, we illustrate the theory of critical slowing down for unidimensional systems more rigorously. To do so, we need to make three key assumptions. First, we assume that the time-series is generated by a dynamical system in equilibrium. Second, we assume that linear stability analysis is informative for the behavior of the system close to the stable equilibrium. Finally, we assume that the system approaches a *zero-eigenvalue* bifurcation.

We first linearize the system at its equilibrium. In linear stability analysis the idea is to simplify the possibly non-linear dynamical system around its equilibrium and map out the dynamics of the simplified system when the state variable is driven out of the equilibrium by some small external perturbation. To model these perturbations explicitly, we rewrite the deterministic one-dimensional system $\dot{x}(t) = f(x(t))$ into its stochastic version:

$$dX(t) = f(X(t))dt + \sigma dW(t) , \quad (3)$$

where $dX(t)$ denotes an Itô integral, σ the diffusion term, and $dW(t)$ is a Wiener process. Equation (3) assumes that every period $[0, T]$ is subject to independent, normally distributed noise of mean zero and variance T . This is the simplest form of introducing noise and other specifications can be considered (see for example O'Regan & Burton, 2018).

Let x^* be a fixed point, i.e. $f(x^*) = 0$, and $\eta(t)$ be a small external perturbation of the system around its stable equilibrium, yielding $X(t) = x^* + \eta(t)$. We linearize $dX(t) = d\eta(t)$ at the stable equilibrium to obtain:

$$d\eta(t) = f(x^* + \eta(t))dt + \sigma dW(t) \quad (4)$$

$$= f(x^*) + f'(x^*)\eta(t)dt + \sigma dW(t) + \mathcal{O}(\eta(t)^2) \quad (5)$$

$$\approx f'(x^*)\eta(t)dt + \sigma dW(t) , \quad (6)$$

since $f(x^*) = 0$ and higher order terms are assumed to be negligible. Equation (4) is an Ornstein-Uhlenbeck process whose solution is (e.g., Gardiner, 2004, p. 106):

$$\eta_T = e^{f'(x^*)T}\eta_0 + \sigma \int_0^T e^{f'(x^*)(T-t)} dW(t) . \quad (7)$$

Recall that a stable equilibrium implies that $f'(x^*) < 0$. In that case, and since the expectation with respect to a Brownian motion is zero, the noise approaches zero in expectation:

$$\lim_{T \rightarrow \infty} \mathbb{E}[\eta_T] = \lim_{T \rightarrow \infty} e^{f'(x^*)T}\eta_0 = 0 . \quad (8)$$

The stationary variance of an Ornstein-Uhlenbeck process for $f'(x^*) < 0$ is given by:

$$\lim_{T \rightarrow \infty} \text{Var}(\eta_T) = -\frac{\sigma^2}{2f'(x^*)} \left(1 - \lim_{T \rightarrow \infty} e^{2f'(x^*)T}\right) = -\frac{\sigma^2}{2f'(x^*)} . \quad (9)$$

This result demonstrates why a system approaching the bifurcation point is expected to exhibit an increase in variance. At the bifurcation point, $f'(x^*)$ changes sign. If the system approaches a bifurcation point from a stable regime, we have $f'(x^*) \rightarrow 0^-$, blowing up the variance to infinity.

A similar result can be shown for the autocorrelation of the time series. The stationary covariance (for finite $\tau = |T - S|$) of the Ornstein-Uhlenbeck process is given by:

$$\lim_{S, T \rightarrow \infty} \text{cov}(\eta_T, \eta_S) = -\frac{\sigma^2}{2f'(x^*)} e^{f'(x^*)\tau} , \quad (10)$$

resulting in an autocorrelation of

$$\lim_{S, T \rightarrow \infty} \text{cor}(\eta_T, \eta_S) = \lim_{S, T \rightarrow \infty} \frac{\text{cov}(\eta_T, \eta_S)}{\sqrt{\text{Var}(\eta_T)\text{Var}(\eta_S)}} = e^{f'(x^*)\tau} . \quad (11)$$

Therefore, if $\lim_{x^* \rightarrow 0^-} f'(x^*)$, the autocorrelation approaches one from below.

A.1 Discrete-time Analysis

Since real-life systems have to be measured at discrete time points, it is instructive to write the result in discrete-time notation. The discrete-time equivalent of the Ornstein-Uhlenbeck process is an autoregressive (AR) model of order one. For the discrete time step $\tau = 1$, the AR(1) can be written as:

$$\eta_t = \eta_{t-1} e^{f'(x^*)} + \varepsilon_t , \quad (12)$$

with $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ and $\sigma_\varepsilon^2 = -\sigma^2 \frac{1 - e^{2f'(x^*)}}{2f'(x^*)}$. The variance of the AR(1) process is then given by:

$$\text{Var}[\eta_t] = \frac{\sigma_\varepsilon^2}{1 - e^{2f'(x^*)}} = -\frac{\sigma^2}{2f'(x^*)} , \quad (13)$$

which is the same as obtain from the continuous-time analysis. Clearly, if $f'(x^*)$ approaches zero from below, the AR(1) approaches the random walk limit with exploding variance.

B Critical Slowing Down for Multidimensional Systems

Here, we illustrate the theory of critical slowing down for multidimensional systems more rigorously. The linearization result of Equation (3) generalizes to the multivariate system:

$$d\eta(t) = J(x^*)\eta(t)dt + SdW(t) , \quad (14)$$

where $\eta(t)$, $dW(t)$, and x^* are vectors of the same length. $J(x^*)$ is the full rank Jacobian matrix evaluated at the fixed point and S a matrix capturing interdependencies of the noise. We again assume the simplest form of independent and normally distributed noise, $S = \sigma\mathbb{I}$. To simplify the exposition, we only consider symmetric Jacobian matrices, as the example considered in the main text.⁷ The result of the multidimensional Ornstein-Uhlenbeck process is derived similarly as in the unidimensional case as:

$$\eta_T = e^{J(x^*)T}\eta_0 + \sigma \int_0^T e^{J(x^*)(T-t)} dW(t) . \quad (15)$$

If the system is in a stable equilibrium, all real parts of the Jacobian's eigenvalues are negative. In that case, the noise wipes out over time and we have:

$$\lim_{T \rightarrow \infty} \mathbb{E}[\eta_T] = \lim_{T \rightarrow \infty} e^{J(x^*)T}\eta_0 = 0 . \quad (16)$$

The stationary covariance for finite $T - S \geq 0$ is given by:

$$\lim_{S, T \rightarrow \infty} \text{cov}(\eta_T, \eta_S^\top) = -\sigma^2 e^{J(x^*)(T-S)} H(x^*)^{-1} , \quad (17)$$

where $H(x^*) = J(x^*) + J(x^*)^\top$. This makes clear that the variance of the system perturbed at the stable equilibrium is given by:

$$\lim_{T \rightarrow \infty} \text{var}(\eta_T) = -\sigma^2 H(x^*)^{-1} . \quad (18)$$

If an eigenvalue of $J(x^*)$ approaches zero from below, the same is true for $H(x^*)$, resulting in the divergence of the variance.

B.1 Discrete-time Analysis

The discrete-time equivalent of the multidimensional Ornstein-Uhlenbeck process is a Vector autoregressive (VAR) process of order one. For discrete time steps of size one, Equation (4) can be written as a VAR(1) model:

$$\eta_t = \Phi\eta_{t-1} + \varepsilon_t , \quad (19)$$

⁷Note that if J is not symmetric, the integral below cannot be solved explicitly.

where $\Phi = e^{J(x^*)}$ and $\varepsilon_t \sim \mathcal{N}(0, -\sigma^2 H(x^*)^{-1})$. The VAR(1) model is stationary if the real parts of the matrix Φ lie within the unit circle, corresponding to the requirement of strictly negative eigenvalues of $J(x^*)$. In that case, the lag-0 variance matrix $\Sigma_0 = \text{Var}(\eta_T)$ becomes:

$$\begin{aligned} \text{Var}(\eta_T) &= \mathbb{E} [\eta_t \eta_t^\top] = \Phi \mathbb{E} [\eta_{t-1} \eta_{t-1}^\top] \Phi^\top + \mathbb{E} [\varepsilon_t \varepsilon_t^\top] \\ \Sigma_0 &= \Phi \Sigma_0 \Phi^\top - \sigma^2 H(x^*)^{-1} . \end{aligned} \quad (20)$$

Equation (20) can be solved explicitly by stacking the columns using the *vec* operator (Hamilton, 1994, p. 265), resulting in:

$$\text{vec}(\Sigma_0) = -\sigma^2 (\mathbb{I} - \Phi \otimes \Phi)^{-1} \text{vec}(H(x^*)^{-1}) , \quad (21)$$

where \otimes denotes the Kronecker product.

Since the eigenvalues of the Kronecker product are just the pairwise products of the eigenvalues of Φ , there exists a finite solution of the covariance matrix. If an eigenvalue of Φ approaches zero from below, Equation (21) diverges, increasing the variance.

From the lag-0 covariance matrix Σ_0 and the VAR(1) coefficient matrix Φ follows the lag- k covariance matrix (Hamilton, 1994, p. 266):

$$\text{cov}(\eta_T, \eta_{T-k}) = \Sigma_k = \mathbb{E} [\eta_t, \eta_{t-k}^\top] = \Phi^k \Sigma_0 , \quad (22)$$

which can be used to make model-based predictions about the strength of the autocorrelation in the individual system components.

C Critical Slowing Down Differs Among Components

We reproduce and extend part of the results by Boerlijst et al. (2013) here to further illustrate the nuances behind critical slowing down. We simulate from the same three-species model with juvenile prey (J), adult prey (A), and a predator (P) which only attacks adult prey:

$$\frac{dJ}{dt} = A - \frac{J^2}{1+J^2} - \mu_J J \quad (23)$$

$$\frac{dA}{dt} = \frac{J^2}{1+J^2} - AP - \mu_A A \quad (24)$$

$$\frac{dP}{dt} = AP - \mu_P P , \quad (25)$$

where (μ_J, μ_A, μ_P) are the death rates of the species. Following Boerlijst et al. (2013), we set $\mu_J = 0.05$ and $\mu_A = 0.10$ and simulate from this model, adding independent noise with $\sigma_\varepsilon = 0.005$ to the death rate of all species.

We slowly varied the death rate of the predators from $\mu_P = 0.45$ to $\mu_P = 0.553$, which is the bifurcation point at which the predators become extinct. We simulated 60,000 time points using a step size of $\Delta = 0.10$ and subsampled the data by taking only every 20th time point; subsampling reduced the magnitude of early warning indicators that are based on correlations, which otherwise would be at ceiling.

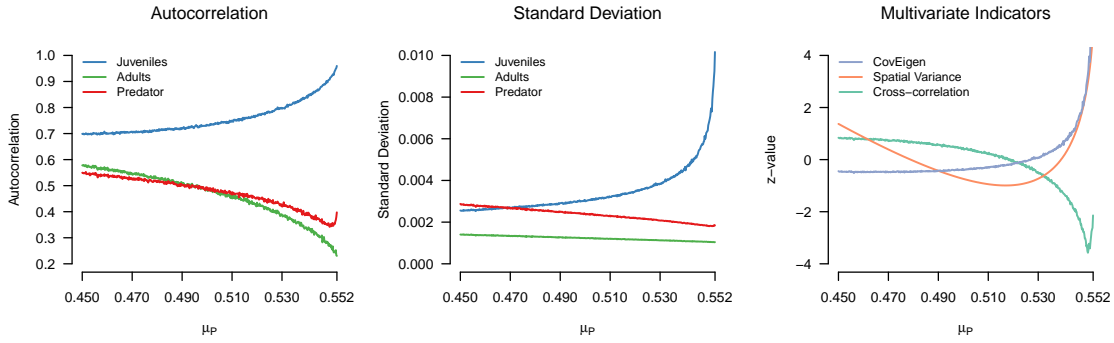


Figure 9: Left: Only juvenile prey showed a continuous increase in autocorrelation as the system approaches the bifurcation point $\mu_P \approx 0.553$. Middle: Only juvenile prey showed an increase in standard deviation. Right: Multivariate indicators (z -standardized) such as the spatial variance and the average absolute value of all cross-correlations first decreased, then increased comparatively slowly (spatial variance) or suddenly (cross-correlation). Only the dominant eigenvalue of the covariance matrix (CovEigen) increased monotonically.

Figure 9 shows that only the juvenile prey showed an increase in autocorrelation as the system approached the bifurcation point $\mu_P \approx 0.553$ (left panel). In fact, the autocorrelation of the adult prey and the predator actually *decreased*, except that there was a rapid increase in autocorrelation for predators right before the bifurcation point.⁸ Similarly, only the juvenile prey showed an increase in standard deviation (middle panel). The right panel in Figure 9 shows that the largest eigenvalue of the covariance matrix — an early warning indicator recently proposed by Chen et al. (2019) — monotonically increased as the system approached the bifurcation point. In contrast, the early warning indicator spatial variance (Guttal & Jayaprakash, 2009; Kéfi et al., 2014), which is defined as the variance of the three-dimensional vector given by the species, and the cross-correlation first *decreased* when approaching the bifurcation point. The spatial variance increased markedly before the bifurcation point, while the cross-correlation increased only immediately before the predators become extinct. The results are qualitatively the same when one adds independent noise not to the death rates (that is, multiplicative noise), but independent noise to the populations (that is, additive noise). Our analysis extends the results presented in Boerlijst et al. (2013), showing that multivariate indicators, too, might be suppressed or show a counter-intuitive decrease when approaching a tipping point. This again underscores the need to have at least partial understanding of the system under study. In particular, while the system might show critical slowing down, it may not be expressed in all system variables, or even the variables that seem most intuitive for critical slowing down to occur. For a rigorous mathematical analysis explaining the conditions under which one can expect early warning signals and in which variables, see Patterson et al. (2021).

D Definition of Early Warning Indicators

Here we provide the mathematical definition of the early warning indicators used in the simulation. Let T denote the number of data points in a particular rolling window. Let x_{ki} denote the i^{th} observation of the k^{th} variable comprising our Generalized Lotka-Volterra model. Let \bar{x}_k denote the sample mean of variable x_k across a rolling window that is apparent from context. The univariate early warning indicators are computed on the variable x_1 and are given by:

⁸Boerlijst et al. (2013) only showed results for $\mu_P \leq 0.55$, which is why they did not observe this sudden increase. Technically, therefore, the predators showed critical slowing down, but it is too little and too late to act in practice.

$$\text{Autocorrelation} = \frac{\sum_{i=1}^T (x_{1i} - \bar{x}_1) (x_{1(i-1)} - \bar{x}_1)}{\sqrt{\sum_{i=1}^T (x_{1i} - \bar{x}_1)^2} \sqrt{\sum_{i=1}^T (x_{1(i-1)} - \bar{x}_1)^2}} \quad (26)$$

$$\text{Variance} = (T - 1)^{-1} \sum_{i=1}^T (x_{1i} - \bar{x}_1)^2 \quad (27)$$

$$\text{Skewness} = \frac{\sum_{i=1}^T (x_{1i} - \bar{x}_1)^3}{\left(\sum_{i=1}^T (x_{1i} - \bar{x}_1)^2\right)^{\frac{3}{2}}} \quad (28)$$

$$\text{Kurtosis} = \frac{\sum_{i=1}^T (x_{1i} - \bar{x}_1)^4}{\left(\sum_{i=1}^T (x_{1i} - \bar{x}_1)^2\right)^2}. \quad (29)$$

Let $\mathbf{x} \in \mathbb{R}^{T \times 4}$ denote the matrix of observations of the four variables for a rolling window of size T . Denote \mathbf{x}_i as the vector of observations of the four variables at time point i , and let $\bar{\mathbf{x}}$ denote the multivariate sample mean over the window of size T . The multivariate indicators are given by:

$$|\text{Cross-Correlation}| = \left| T^{-2} \sum_{k=1}^T \sum_{l=1}^T \frac{\sum_{i=1}^T (x_{ki} - \bar{x}_k) (x_{li} - \bar{x}_l)}{\sqrt{\sum_{i=1}^T (x_{ki} - \bar{x}_k)^2} \sqrt{\sum_{i=1}^T (x_{li} - \bar{x}_l)^2}} \right| \quad (30)$$

$$\text{Spatial-Variance} = T^{-1} \sum_{i=1}^T (\mathbf{x}_i - \bar{\mathbf{x}})^2 \quad (31)$$

$$\text{Spatial-Skewness} = \frac{\sum_{i=1}^T (\mathbf{x}_i - \bar{\mathbf{x}})^2}{\left(\sum_{i=1}^T (\mathbf{x}_i - \bar{\mathbf{x}})^2\right)^{\frac{3}{2}}} \quad (32)$$

$$\text{Spatial-Kurtosis} = \frac{\sum_{i=1}^T (\mathbf{x}_i - \bar{\mathbf{x}})^4}{\left(\sum_{i=1}^T (\mathbf{x}_i - \bar{\mathbf{x}})^2\right)^2}, \quad (33)$$

where the exponents indicate element-wise operations. The dominant eigenvalue of the covariance matrix is the largest absolute eigenvalue of the covariance matrix, see Chen et al. (2019), for details.

E Further Simulation Results

Figure 10 shows the average Area under the Curve (AUC) across settings for the remaining early warning indicators.

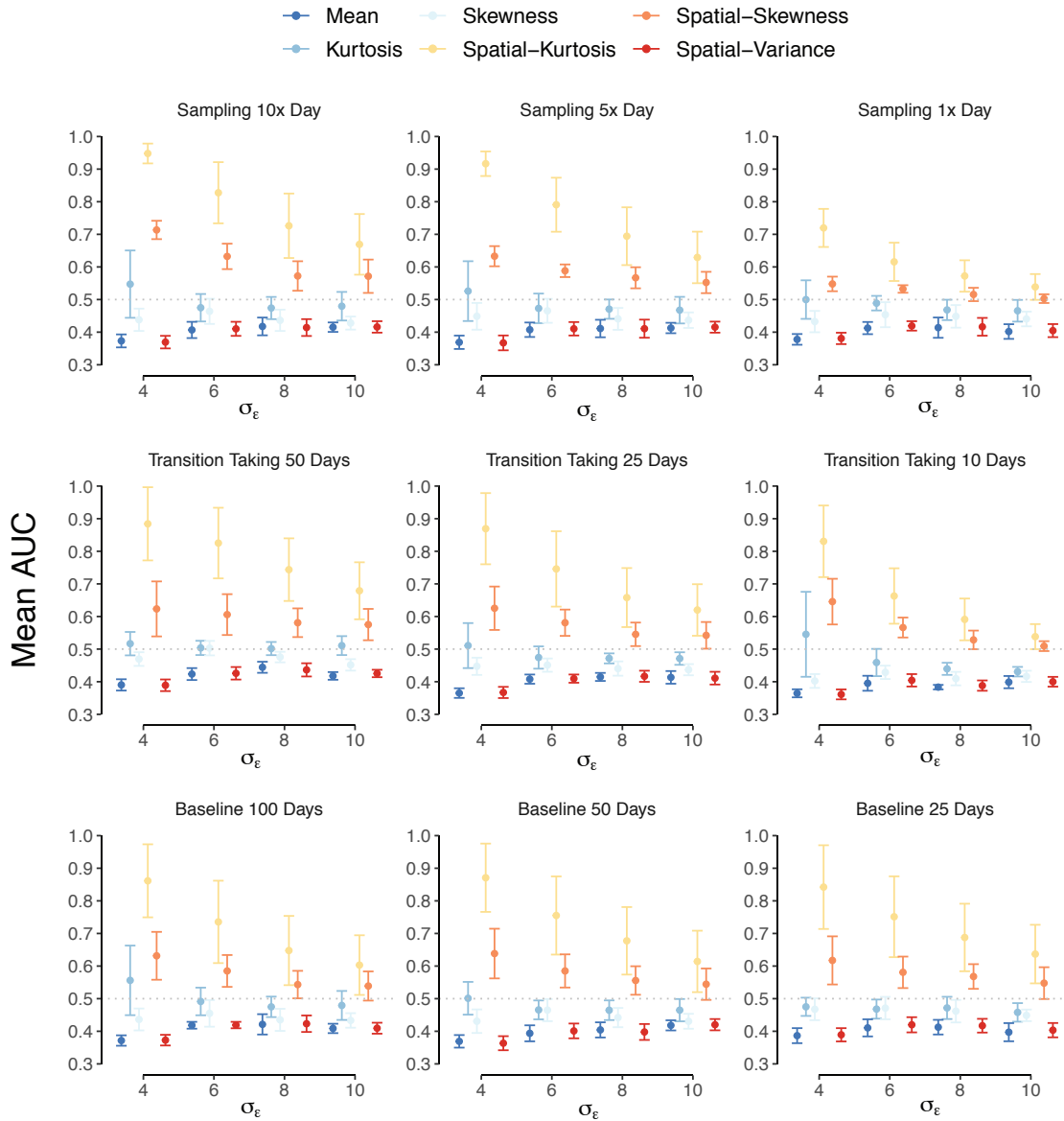


Figure 10: Shows Area under the Curve (AUC) of selected early warning indicators averaged over the transitioning period across sampling frequencies (top panels) and averaged over sampling frequencies across transitioning period (bottom panels). All results are further averaged over baseline (25, 50, 100 days), and rolling window (10, 25, 50 days); error bars denote one standard deviation across these configurations.