CESifo WORKING PAPERS

10892
2024

January 2024

Balancing the Risk of Tipping: Early Warning Systems from Detection to Management

Florian Diekert, Daniel Heyen, Frikk Nesje, Soheil Shayegh

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 10892*

Balancing the Risk of Tipping: Early Warning Systems from Detection to Management

Abstract

Early warning signals (EWS) of imminent regime shifts can be identified through the observation of a system's behavior under increasing stress and before crossing a tipping point. Despite many advances in the detection of EWS in recent years, EWS are yet to find direct application in management. Here, we focus on operationalizing the EWS information in an early warning system consisting of a tipping indicator (e.g., autocorrelation), whose value increases as the system approaches the tipping point, and a trigger value, beyond which an EWS is sent. We demonstrate how such an early warning system allows managers to balance the risk of tipping by providing information for updating their belief about the location of the tipping point. In particular, deployment of an early warning system results in taking more cautious early steps while it encourages more risk taking behavior in later stages if no EWS is sent. We uncover a tension between better information about the location of the tipping point and increased risk of crossing it as a result of EWS. Our framework complements the emerging EWS knowledge in the natural sciences with a better understanding of how, when, and why EWS improve management.

JEL-Codes: C610, D830, Q540.

Keywords: catastrophic regime shifts, tipping points, early warning signals, learning, optimal ecosystem management.

*Florian Diekert**
Centre for Climate Resilience
University of Augsburg / Germany
florian.diekert@uni-a.de

Daniel Heyen
Department of Economics
RPTU Kaiserslautern-Landau / Germany
daniel.heyen@wiwi.rptu.de

Frikk Nesje
Department of Economics
University of Copenhagen / Denmark
frikk.nesje@econ.ku.dk

Soheil Shayegh
RFF-CMCC European Institute on Economics and the Environment (EIEE), Milan / Italy
soheil.shayegh@eiee.org

*corresponding author

We thank Lassi Ahlvik, Stephen R. Carpenter, Eric Navdal, and Andries Richter for discussions. The research has been funded by the European Research Council Project NATCOOP (ERC StGr 678049).

1 Introduction

Human activities push many ecosystems over their tipping points where large, rapid, and often irreversible changes are expected to follow. Indeed, for some systems, such as the global climate (Boers, 2021; Cai et al., 2015; Lenton et al., 2019; Dietz et al., 2021; McKay et al., 2022) and the Amazonian rainforest (Hirota et al., 2011; Lovejoy and Nobre, 2018), the sheer scale and irreversibility of potential damages make it imperative to predict the tipping points and act ahead to prevent crossing them.

A growing literature in the natural sciences is dedicated to the identification of generic early warning signals (EWS) of tipping points (Scheffer et al., 2009). EWS can be detected in a variety of natural systems from lake ecosystems (Carpenter et al., 2011), forests (Liu et al., 2019; Boulton et al., 2022) and fisheries (Clements et al., 2017) to climate physics (Boers, 2021; Ditlevsen and Ditlevsen, 2023). The detection of EWS has also been studied in many social fields and applications, ranging from medical sciences (van de Leemput et al., 2014; Helmich et al., 2022) and psychology (Hart et al., 2020) to epidemiology (O'Brien and Clements, 2021), finance (Sarkar and Sriram, 2001; Wen et al., 2018), land degradation (Bruzzone and Easdale, 2021), and urban planning (Dianat et al., 2022). However, what is largely overlooked by this literature is how to actually operationalize EWS and put them to best use for sustainable resource management. This paper takes a first step in this direction by developing a framework to address the question of how to build effective early warning systems (EWSys) to improve the management of at-risk systems.

Our proposed framework, shown in Figure 1, builds on and combines the EWS detection literature in the natural sciences with the economic literature on optimal management under tipping risk. We connect these two literatures to demonstrate the necessary steps for building effective EWSys for management.

Note that detecting EWS is different from forecasting the system's behaviour. The EWS literature uses the idea that the time series of the observations of a socio-ecological system's behaviour contains a signature that can anticipate tipping (Dakos et al., 2015; Bury et al., 2021; Dakos et al., 2023). This is fundamentally different from forecasting which is based on collecting and combining co-variates to predict the system's state in the future (Petropoulos et al., 2022). Thus, the central feature of EWS detection is that it does not need training or calibration. Therefore, EWS can be applied to systems for which replicates are not available (such as the planet's climate, the Amazon, or other large ecosystems, Scheffer et al. (2009); Boettiger and Hastings (2012)). The other aspect that distinguishes EWS detection is that unlike forecasting, it can be applied to problems where the tipping risk arises endogenously, i.e., when the system's behaviour depends on or is influenced by controllable human actions. This feature links EWS detection to the fields of economics and management where controllable actions take center stage. This includes important contributions to better understanding of how to react to endogenous or exogenous tipping threats (Polasky et al., 2011; Crépin and Nævdal, 2019), and how to manage specific socio-ecological systems under tipping risk ranging from groundwater management to the management of fisheries and climate change (Tsur and Zemel, 1995; Nævdal, 2006; Cai and Lontzek, 2019; Voss and Quaas, 2022), as well as first advances to incorporate Bayesian learning (Diekert, 2017; Lemoine and Traeger, 2014). Our paper squarely fits in this interdisciplinary domain by introducing Bayesian belief revision (Sarkar and Sriram, 2001) to an optimal management problem.

Managing real socio-ecological systems involves many actors such as politicians, bureaucrats, scientists, resource users and interest groups. Here, we differentiate and focus on only two entities within the management realm: the *regulator* that makes decisions on resource use, and the *scientific agency* which monitors and reports the state of the socio-ecological system to the regulator (see Figure 1). In other words, the scientific agency detects the tipping point while the regulator makes decisions about how to best respond to it.

Specifically, our stylized framework for operationalizing early warning signals considers a regulator who needs to decide how to manage a socio-ecological system that has an unknown and constant tipping point. The regulator's objective is to maximise yield from the system without pushing it over the tipping point. However, once pressure exceeds a critical value, the system collapses. The scientific agency monitors the stock status and uses the data to calculate a tipping indicator. The tipping indicator and trigger value together constitute an EWSys. If the tipping indicator exceeds the trigger value, a signal (EWS) is sent to the

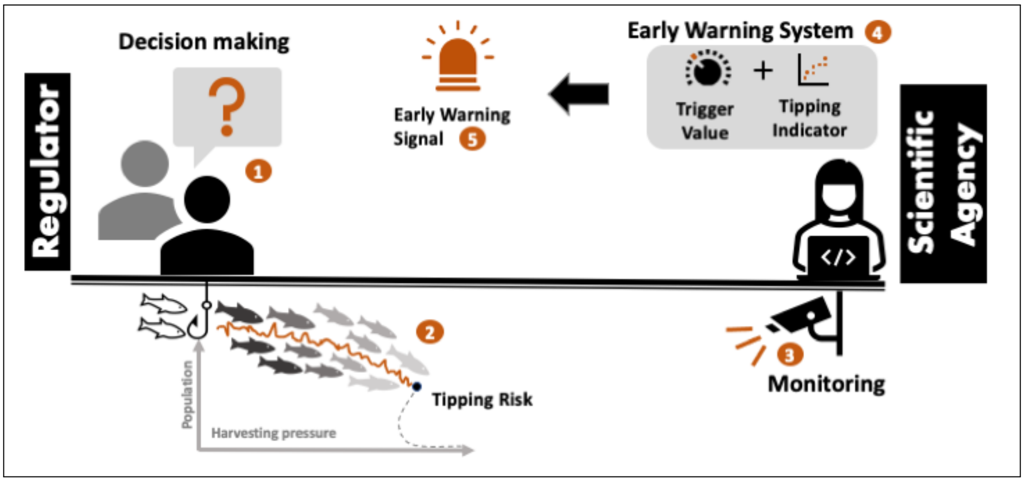


Figure 1: Operationalizing early warning signals: The regulator needs to decide how much to harvest (1) from a socio-ecological system that has an unknown tipping point (2). The scientific agency monitors the stock development (3) and uses the data to calculate a tipping indicator. The tipping indicator and trigger value together constitute an early warning system (4). If the tipping indicator exceeds the trigger value, an early warning signal (EWS) is sent to the regulator (5). Based on the information received, the regulator then decides whether to increase harvesting pressure (1).

regulator. Based on the information received, the regulator accordingly updates her knowledge about the location of the tipping point and then decides whether to increase harvesting pressure.

The management process illustrated in Figure 1 is generic and could also apply to other socio-ecological systems such as land-use decisions or timber extraction from a rainforest whose self-sustaining moisture system collapses, once the size of the remaining forest is too small. It could equally well apply to managing groundwater extraction from an aquifer under the threat of saltwater intrusion once the water table falls below a critical value, or to controlling an animal population, such as a fish stock, that collapses once its size falls below a minimum viable threshold. For concreteness, we consider the latter example as our leading case.

Our innovation is to shift the focus from detection to action and show how and why early warning systems improve management, under what conditions they will be more effective, and what potential risks they may entail. As such, our paper makes three main contributions. The first key contribution is to build a better understanding of the effect of EWSys on management through a step-by-step process, starting from the simulation of a socio-ecological time series from the natural sciences literature (Bury et al., 2020; Dakos et al., 2023) and ending with the optimal sequence of decisions. The proposed framework, shown in Figure 1, is quite general and can be applied to many other socio-ecological systems beyond what is discussed in this paper. Our second key contribution is to spell out the conditions under which different parameters, such as the regulator's time preference and risk tolerance, affect the optimal design and management implication of an early warning system in this framework. A related question is how the scientific agency should set the trigger value. As our framework paves the way for fruitful collaboration across different disciplines, we emphasize the potential benefits and risks of relying on the EWS in concrete management applications ranging from ecosystems and sustainable resource use to emergency planning. In particular, the third key contribution of the paper is to uncover and highlight an underlying tension between the value of better information and risk: while the optimal use of EWSys improves ecosystem management, it can lead to an increase in the risk of tipping.

Additionally, our framework points to important avenues for further research on EWS applications in management science. For example, it is important to consider how real world decision makers would react to EWS under impending regime shift risks (Seifert et al., 2023). Another example is how the availability of an EWSys would impact management outcomes when the regulator and scientific agency can act strategically (Alizamir et al., 2020).

The rest of the paper proceeds as follows. We present our framework in Section 2. Concretely, we outline model components such as the definition of an early warning system based on the underlying socio-ecological system and trigger value, as well as the specification the management problem and main outcomes of interest. Next, in Section 3 we characterize the optimal solution, with particular focus on tipping risk and economic value, and detail important dimensions of EWSys design. Finally, we offer a concluding discussion in Section 4.

2 The framework

We start with a model of a socio-ecological system that has a unique but unknown tipping point a_{crit}. The tipping point can be anywhere on the interval $[0, \bar{a}]$. The system can accommodate some pressure from human actions a and still maintain its sustainability. However, once pressure exceeds a critical value, that is $a > a_{crit}$, the system collapses. In plain words, a collapse of the system occurs whenever the regulator's action exceeds the tipping point. We couple this model with a two-period dynamic management setup. That is, the regulator chooses an action for each period. After deciding the first period action, and before deciding the second period, the regulator may receive an *early warning signal* (EWS) from the scientific agency. While our framework is stylized to isolate how *early warning systems* (EWSys) affect optimal decisions, it is applicable to complex socio-ecological systems.

For concreteness, we consider a fishery as a leading example of a socio-ecological system under the threat of collapse due to increased pressure from human activities (i.e., over-fishing). The regulator decides how much fish should be harvested in a given period by setting a quota, allocating it to individual harvesters, and enforcing compliance. The scientific agency, on the other hand, gathers information on the state of the fish stock (e.g., through stock assessments, trawl surveys, oceanographic research) and informs the regulator if the fish population is nearing the tipping point.[1]

In our model, the regulator's objective is to maximise harvest without pushing the fishery's population over the tipping point. Further, we assume that the regulator takes action in two consecutive stages (period 1 and period 2) with a time preference for the first period. The variables a_1 and a_2 denote the action that the regulator takes in period 1 and period 2, respectively, while $u(a)$ is the utility, derived from a given action a, normalized between zero and one by dividing it by the highest possible harvesting action $\bar{a}$. The time preference manifests itself

[1] In some countries, these two dimensions are clearly separated (in Norway, for example, the *Directorate of Fisheries* administers fishing regulations, while the *Institute for Marine Science* provides stock assessments and other scientific advice). In other countries the two management dimensions are represented by the same agency (in the US, for example, *NOAA* is responsible for both research and regulation).

in the form of a discount factor, $\beta \in (0, 1]$, which will play an important role in determining optimal actions under uncertainty about the tipping point.[2]

We assume that the regulator holds a uniform prior belief for the tipping point on $[0, \bar{a}]$ and denote by $P_1(a_1)$ the probability that the action a_1 does not cause the collapse of the socio-ecological system. Crossing the tipping point triggers an irreversible collapse of the population and brings zero value to the regulator, the same as no harvesting, $a = 0$. Hence, the collapse of the resource could be interpreted as the extinction of the resource stock, or as a moratorium on harvesting. While this is a convenient normalization, the essential feature for our analysis is that the post-collapse utility does not depend on pre-collapse actions, see e.g., Diekert (2017).

Facing this risky prospect implies that the regulator's risk tolerance will also play an important role in determining optimal actions. The regulator's risk tolerance is given by the parameter $\alpha \in (0, 1]$, which follows from assuming that $u(a) = (a/\bar{a})^{\alpha}$. On top of that, there is also learning. There is a probability function $P_2(a_2)$ that describes whether the system remains stable after taking a given action in the second period. It differs from $P_1(a_1)$ in the sense that choosing $a_1 > 0$ and not causing collapse means that the tipping point was not below a_1. Hence, after period 1, the regulator learns that $a_1 \leq a_{crit}$ (and hence $a_{crit} \in [a_1, \bar{a}]$).

In addition, the regulator may receive an EWS from the scientific agency about the state of the fishery, if there exists an early warning system. Receiving or not receiving an EWS is the second source of information that affects $P_2(a_2)$. Accordingly, the regulator updates her prior belief about the location of the tipping point before deciding on the action in the second period, a_2. The management problem is therefore to find the sequence of actions $\{a_1, a_2\}$ which yields maximum harvest without causing the system to collapse. We argue that to solve this problem, managers can use an early warning system to increase their gain, but we also caution against relying on EWS for reducing the risk of collapse.

We now describe how such an early warning system look like.

[2] Note that we assume that tipping is "nonstochastic". This reflects that the tipping point is an imminent characteristic of the system. Any stochasticity would "serve to approximate a more complete model with uncertainty [...] over the precise trigger mechanism underlying the tipping point" (Lemoine and Traeger, 2014, p.155), and not be a genuine feature of tipping.

2.1 Designing an Early Warning System

2.1.1 The socio-ecological system

First, we describe the underlying model governing the socio-ecological system. To appreciate this model, it is important to note that we distinguish two time scales. The regulator acts twice, once in period 1 and once in period 2. The underlying resource develops over many short time steps t in between. In terms of our concrete leading example, one can think of the periods as years and the regulator sets a harvest quota for year, while the time steps are days.

For simplicity, we assume, in line with other workhorse models (Bury et al., 2021; Boers, 2021), that population growth subject to harvesting takes a quadratic form:

$$g(x_t, a_t) = x_{t+1} - x_t = -\gamma_2\ x_t^2 + \gamma_1\ x_t - a_t + \sigma\epsilon_t, \tag{1}$$

where x_t is the population size at time t, γ_1 and γ_2 are the linear and quadratic growth parameters, a_t is the harvesting action, ϵ_t is a white noise error term that follows a standard normal distribution and σ is the noise amplitude.

In the absence of harvesting, the carrying capacity of this ecosystem is at $K = {}^{\gamma_1}/_{\gamma_2}$ from $\mathrm{E}[g(x_t, 0)] = 0$ where E is the expectation operator. Furthermore, the maximum population growth is achieved when $x^* = {}^{\gamma_1}/_{2\gamma_2}$. This provides a basis for defining the highest action that keeps the ecosystem at a sustainable level:

$$E[g(x^*, a_{crit})] = 0, \text{ and from (1) we get } a_{crit} = {}^{\gamma_1^2}/_{4\gamma_2}. \tag{2}$$

In our stylized setting, this critical action represents the location of the tipping point a_{crit}. If the regulator increases the action from its initial value of zero, the population is able to recover and sustain its reproductive level as long as the action does not exceed a_{crit}. Once the action exceeds the critical value a_{crit}, the population will collapse and will not be able to recover to its sustainable level.

Figure 2a shows one realization of our population model with an unknown tipping point while facing increasing harvesting pressure. When generating it, we set $\sigma = 0.1$ and $\bar{a} = 2$. Figure 2a is produced by increasing harvesting action along the horizontal axis (grid of 500 points, $a \in \{0, 0.004, \ldots, 2\}$) in increments of 0.004 per time step for a population model. A model is defined by its linear

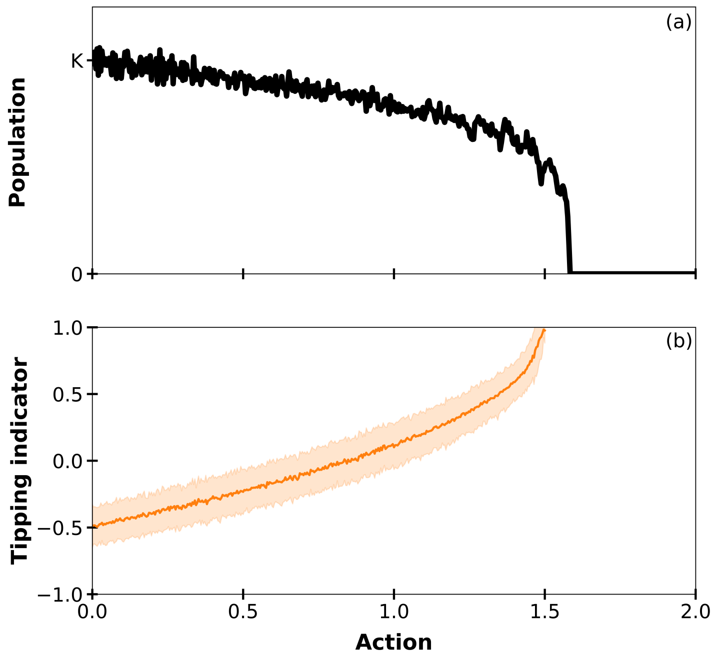


Figure 2: (a) Population dynamic for a fishery with an unknown tipping point subject to an increasing harvesting action. The initial population is indicated by K. (b) The tipping indicator (e.g., autocorrelation) increases as the ecosystem approaches the tipping point.

and quadratic growth parameters γ_1 and γ_2. As we set the carrying capacity $K = \gamma_1/\gamma_2 = 4$, we have $\gamma_1 = 4a_{crit}/K = a_{crit}$ and $\gamma_2 = a_{crit}/4$ so that any model can be identified uniquely by the value of a_{crit}. Therefore, uncertainty over parameters of the socio-ecological system means uncertainty over the tipping point location a_{crit}. In the management problem, the scientific agency uses data from the EWSys to infer whether tipping is likely to be induced by increasing action beyond a_1.

2.1.2 Tipping indicators

Although the exact location of the tipping point is unknown to the regulator, a literature in the natural sciences (Scheffer et al., 2009, see also Dakos et al., 2023 for a recent review) has identified some statistical characteristics of the underlying dynamic system that can act as a *tipping indicator*.[3] Following this literature, we apply, for each action a along the horizontal axis (grid of 500 points, $a \in \{0, 0.004, \ldots, 2\}$ with $n = m = 100$), the following algorithm to generate Figure 2b:

1. calculate the steady-state population level corresponding to action a,
2. draw n random samples from the normally distributed noise parameter ϵ_t to produce a time series of the population, starting from the steady-state population, when action a is applied,
3. calculate the autocorrelation of lag 1 for this time series and report it as the tipping indicator,
4. repeat this procedure for m times to obtain confidence intervals for the tipping indicators (shaded areas in Figure 2b).

The algorithm connects the action in the first period with the action in the second period of the management problem: the first period action induces stock variations.[4] Assuming that the first period action did not lead to collapse, there is sufficient time for the system to get to a stable steady state, and for calculating the tipping indicator, before the second period action is taken. After observing the state of the system and its tipping indicator, the scientific agency indicates an EWS or no EWS and sends it to the regulator who then decides on the second period action. In terms of the leading example, one could think that the action in each period represents the harvest quota that is set once a year while each time step in the population model, t, represents a day in which a sample of the population size is taken. Alternatively, the periods could represent decades and the actions represent long-term management plans or harvest control rules that are

[3]The choice of the tipping indicator depends on the characteristics of the socio-ecological system. Dakos et al. (2023) have identified 65 different tipping indicators that have been used in the literature. The most commonly used are autocorrelation, skewness, and power spectrum (Held and Kleinen, 2004; Dakos et al., 2012; Clements et al., 2015; Bury et al., 2020)).

[4]Note that this algorithm does not allow for learning about the tipping point from the transition to the steady state. Generally, there could be learning from (i) the system's approach to steady state, (ii) whether or not the system tips, and (iii) the behaviour of the tipping indicator. Our focus in this paper is on the two latter cases. This is in line with the EWS literature (Dakos et al., 2008; Scheffer et al., 2009; Boettiger and Hastings, 2012; Dakos et al., 2015; Bury et al., 2020, 2021; Boers, 2021). We leave the analysis of how the regulator can learn from the approaching path for future research.

revised periodically. The important point is that the scientific agency has access to a time series of observations that comes at a much higher frequency than the decisions about the actions that the regulator takes.

2.1.3 Trigger values and signal functions

As discussed extensively in the tipping point detection literature (Dakos et al., 2015), the tipping indicator behaves in a way that its value can be used as a proxy for the risk of crossing the tipping point. For example, auto-correlation shown in Figure 2b increases as the system nears tipping. Furthermore, to design an early warning system (EWSys), one can compare the value of the tipping indicator against a benchmark called *trigger value*. When the value of the tipping indicator exceeds the trigger value, an alarm (also known as EWS) goes off (Figure 1) signaling that the system is within the vicinity of the tipping point. The choice of trigger value in this sense, determines the sensitivity and specificity of the EWSys. Therefore, we define an EWSys as the combination of a tipping indicator and a trigger value.

However and because the indicator is noisy, we can describe the indicator as a cumulative distribution function F, where $F(a, a_{crit}, \theta)$ is the probability that the indicator does not exceed the value θ if the current action is a and the tipping point is at a_{crit}.

We now introduce the *signal function* $f_{a_{crit},\theta}(d)$ that describes the probability of an EWS when the distance to the tipping point is $d = a_{crit} - a$. Beyond not tipping, it is by receiving or not receiving the EWS that the regulator learns whether the tipping point of the socio-ecological system is likely to be near or far away. For a given trigger value θ, we say that the scientific agency reports an EWS if the tipping indicator exceeds the trigger value. The definition of F as a cumulative distribution function representing the indicator distribution implies that the probability of an EWS (i.e., that the tipping indicator exceeds the trigger value θ) is

$$f_{a_{crit},\theta}(d) = 1 - F(a_{crit} - d, a_{crit}, \theta). \tag{3}$$

Because F is a cumulative distribution function, it follows that $f_{a_{crit},\theta}(d)$ decreases in the trigger value θ for all a_{crit} and d.

Depending on the context we may also want to write the signal function for a fixed action a (instead of fixing tipping point a_{crit}). We then get

$$\tilde{f}_{a,\theta}(d) = 1 - F(a, a + d, \theta). \tag{4}$$

For some concrete applications, the tipping indicator may give rise to signal functions that only depend on the relative distance d and not on the tipping point (or equivalently, the action), $f_{a_{crit},\theta}(\cdot) = f_{a'_{crit},\theta}(\cdot)$ for all a_{crit}, a'_{crit}. The theoretical analysis of such cases is more tractable. In this paper however, we do not make any simplifying assumption of that sort and work with the general forms (3) and (4).

We generate Figure 3, taking as input the socio-ecological system – including the action space (grid of 500 points, $a \in \{0, 0.004, \ldots, 2\}$) and tipping point space (grid of 500 points, $a_{crit} \in \{0, 0.004, \ldots, 2\}$). For autocorrelation, we consider trigger values between -1 and 1 in increments of 0.004 (grid of 501 points, $\theta \in \{-1, -0.996, \ldots, 1\}$). The translation from Figure 2b to Figure 3 works as described in equations (3) and (4) when applying specific trigger values θ.[5]

This captures the notion that an EWSys with a given tipping indicator and a low trigger value sends EWS more often than an EWSys with the same tipping indicator but with a higher trigger value. Similarly, for an EWSys with a given trigger value, the closer the population is to collapsing, the higher is the value of the tipping indicator, and thus the more likely it is that the scientific agency reports an EWS to the regulator.[6]

While Figure 3a illustrates the generic shape of the signal function, Figure 3b shows the family of signal functions that result from a range of different trigger values $\theta \in [-1, 1]$. In other words, Figure 3b translates the data shown in Figure 2b into the probability of receiving an EWS which mimics the probability of tipping

[5]For clarity, in the management problem we work with a smooth version of the signal function. The smooth version is a moving average of the original signal functions. In a first step, we calculate a moving average of span 13 (6 points to each side) of the first signal function component (current location), then the second (actual tipping point), then the third (trigger value). We repeat this procedure twice with span 9 (4 points to each side). Also see Figure A1.

[6]Compared to the EWS literature that considers early warnings to be a function of exogenous time and then spells out sufficient time to act (e.g. through the formulation of "lead time"), EWS in our paper depend on the endogenous distance to the tipping point in terms of harvesting pressure. Hence, the notion of EWS is directly relevant for ecosystem management.

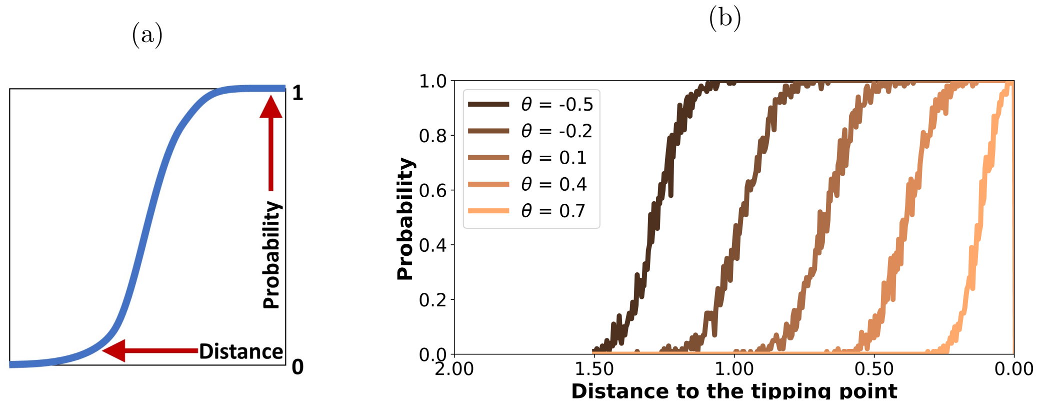


Figure 3: (a) The signal function expresses the probability of receiving EWS as a function of distance to the tipping point. (b) A family of signal functions can be derived for a given tipping indicator (here: autocorrelation, Figure 2b) and a set of trigger values (here ranging from $\theta = -0.5$ to θ=0.7). When the distance to the tipping point is 0.5 units, the probability to receive an EWS is close to 0% for an EWSys with a trigger value of θ=0.7, about 2% for an EWSys with a trigger value of θ=0.4, about 94% for an EWSys with a trigger value of θ=0.1, and nearly 100% for an EWSys with a trigger values of $\theta = -0.2$ and $\theta = -0.5$.

itself.

In our leading example, the regulator sets a given quota for the first period a_1 knowing that EWSys will warn her if this quota would push the system close to the tipping point a_{crit}. If the first-period quota did not lead to collapse, she will receive an update from the scientific agency whether there is an EWS or not. Receiving or not receiving an EWS is helpful information for setting the second-period quota a_2. If there is no EWS, the regulator realizes that increasing the harvesting action (at least, by a small bit) is unlikely to cause the collapse of the fish stock in the second period. But if there is an EWS, the regulator realizes that the current harvesting quota has pushed the system close to the tipping point. Therefore, increasing the harvesting action is not wise, as it may cause the system to tip. To put it in technical terms, receiving or not receiving an EWS helps the regulator to update her belief about the location of the tipping point. The Appendix illustrates the concept of Bayesian learning for two different EWSys with low and high trigger values θ and with low and high levels of harvesting actions in the first period a_1, considering the resulting family of signal functions $f_{a_{crit},\theta}$ as input. We now turn our attention to how to utilise EWS to design an optimal strategy to improve management outcomes.

2.2 The management problem

We first detail the regulator's objective function. In general, the value of the two-period management problem is given by (5):

$$V_1 = \max_{a_1 \in [0,\bar{a}]} \Big\{ P_1(a_1) \cdot \Big(u(a_1) \ + \ \beta \Big[q(a_1) \cdot V_2^{EWS}(a_1) + (1 - q(a_1)) \cdot V_2^{noEWS}(a_1) \Big] \Big) \Big\}, \tag{5}$$

$$\begin{aligned} \text{where} \quad V_2^{EWS} &= \max_{a_2 \in [0,\bar{a}]} \left\{ P_2^{EWS}(a_2) \cdot u(a_2) \right\} && \text{if EWS received} \\ V_2^{noEWS} &= \max_{a_2 \in [0,\bar{a}]} \left\{ P_2^{noEWS}(a_2) \cdot u(a_2) \right\} && \text{if no EWS received} \end{aligned}$$

The connection between utility, actions, and the socio-ecological model in equation (1) comes through $P_1(a_1) = 1 - \int_0^{a_1} p_1(a_{crit}) da_{crit}$, the probability that a given action a_1 does not cause the system to collapse. That probability depends on the action (a higher action increases the risk of collapse and accordingly reduces P) and the probability distribution p_1, the regulator's belief about the tipping point location at that time. We assume that the regulator holds a uniform prior belief for the tipping point on $[0, \bar{a}]$.

This belief is updated between the first and the second period based on two sources of information: First, from experience as she learns whether the first period action caused the system to tip or not. Second, from the EWSys (if it exists) where she learns whether an EWS is received or not. Her second period belief will be either

$$p_2^{EWS}(a_{crit}) = \frac{f_{a_{crit},\theta}(a_{crit} - a_1) p_1(a_{crit})}{\int_{a_1}^{\bar{a}} f_{\tilde{a}_{crit},\theta}(\tilde{a}_{crit} - a_1) p_1(\tilde{a}_{crit}) d\tilde{a}_{crit}} \tag{6}$$

or

$$p_2^{noEWS}(a_{crit}) = \frac{(1 - f_{a_{crit},\theta}(a_{crit} - a_1)) p_1(a_{crit})}{\int_{a_1}^{\bar{a}} (1 - f_{\tilde{a}_{crit},\theta}(\tilde{a}_{crit} - a_1)) p_1(\tilde{a}_{crit}) d\tilde{a}_{crit}}, \tag{7}$$

where $f_{a_{crit},\theta}$ is the signal function.

Because the regulator anticipates that her beliefs will change after the first period action, problem (5) has a recursive structure. The probability to receive an EWS is given by

$$q(a_1) = \frac{\int_{a_1}^{\bar{a}} f_{a_{crit},\theta}(a_{crit} - a_1) p_1(a_{crit}) da_{crit}}{P_1(a_1)}. \tag{8}$$

We will refer to q as the EWS probability. It depends on the action in period 1, the initial belief p_1 about the true location of the tipping point, and the signal function $f_{a_{crit},\theta}$ that describes the likelihood of receiving an EWS when located at a distance d to the true tipping point a_{crit}. As described above, the trigger value θ determines the shape of that signal function. The denominator $P_1(a_1)$ reflects that the correct interpretation of q is EWS probability, provided that action a_1 did not lead to collapse.

2.3 The effect of an Early Warning System

In addition to characterizing the optimal solution with an early warning system, we compare the situation with an EWSys to the situation without an EWSys. Here, we concentrate on two outcomes: The risk of tipping, and the value that a given EWSys provides over a situation without an EWSys. We will evaluate how the risk of tipping and the added-value of the EWSys depend on the trigger value, and the regulator's time and risk preferences.

To describe the change in tipping risk and the value that a given EWSys provides over the baseline situation without an EWSys, we need to define the latter. We call the value of (5) when no EWSys is available V^b and the corresponding tipping risk R^b.

The probability of not crossing the tipping point when choosing a harvesting action $a \in [0, \bar{a}]$ is $P(a) = \frac{\bar{a}-a}{\bar{a}}$. Building on the fact that the first and second period actions are identical in the model without an EWSys (Diekert, 2017), and inserting the utility function $u(a) = (a/\bar{a})^\alpha$, we can express the regulator's objective function as:

$$V^b = \max_a \left\{ \frac{\bar{a} - a}{\bar{a}} (1 + \beta)(a/\bar{a})^\alpha \right\}. \tag{9}$$

The first-order condition for this problem is

$$-a^{\alpha} + (\bar{a} - a)\alpha a^{\alpha-1} = 0, \tag{10}$$

which yields the optimal baseline action:

$$a^b = \frac{\alpha}{1+\alpha}\bar{a}. \tag{11}$$

2.3.1 Definition of tipping risk

With an EWSys, we will show that it is generally not optimal to choose a^b in both the first and the second period. Rather, optimal actions will generally differ between the first and the second period and whether an EWS has been received or not. Given an optimal contingent plan $\{a_1^*,\, a_{2,EWS}^*,\, a_{2,noEWS}^*\}$, we determine the probability that these actions will lead to a collapse of the system, the tipping risk. We assume that the regulator's prior belief about the location of the tipping point is unbiased, i.e., the true location of the tipping point is the realization of a random draw from a distribution that coincides with the regulator's prior belief.

Consider one tipping point a_{crit}. Using the indicator function $\mathbb{1}_{(Cond)}$ that equals 1 if condition $Cond$ is true and 0 otherwise, we can write the condition for collapse as

$$C(a_{crit}) = \mathbb{1}_{(a_1^*>a_{crit})} + \mathbb{1}_{(a_1^*\leq a_{crit})}\cdot\left[\mathbb{1}_{(EWS)}\cdot\mathbb{1}_{(a_{2,EWS}^*>a_{crit})} + \mathbb{1}_{(noEWS)}\cdot\mathbb{1}_{(a_{2,noEWS}^*>a_{crit})}\right]. \tag{12}$$

One possibility for collapse is that the first period action already exceeds the tipping point (the first part in (12)). If the first action does not exceed the tipping point, then tipping occurs with an EWS, if the action $a_{2,EWS}^*$ exceeds the tipping point, and without an EWS, if the action $a_{2,noEWS}^*$ exceeds the tipping point.

We can now calculate the total probability of collapse as $R = \int C(a_{crit})p_1(a_{crit})da_{crit}$, where p_1 is the prior. We get

$$R = \int_0^{a_1^*} p_1(a_{crit})da_{crit} + \int_{a_1^*}^{a_{2,EWS}^*} f_{a_{crit},\theta}(a_{crit} - a_1^*)p_1(a_{crit})da_{crit} + \int_{a_1^*}^{a_{2,noEWS}^*} (1 - f_{a_{crit},\theta}(a_{crit} - a_1^*))p_1(a_{crit})da_{crit}. \quad (13)$$

The baseline tipping risk that is induced by the optimal action a^b in the absence of an EWSys is $R^b = \int_0^{a^b} p_1(a_{crit})da_{crit}$. Below we establish that $a_1^* \leq a^b$ and $a_1^* \leq a_{2,EWS}^* \leq a_{2,noEWS}^*$. Because of this, we can write the change in risk due to the EWSys as

$$R - R^b = -\int_{a_1^*}^{a^b} p_1(a_{crit})da_{crit} + \int_{a_1^*}^{a_{2,EWS}^*} f_{a_{crit},\theta}(a_{crit} - a_1^*)p_1(a_{crit})da_{crit} + \int_{a_1^*}^{a_{2,noEWS}^*} (1 - f_{a_{crit},\theta}(a_{crit} - a_1^*))p_1(a_{crit})da_{crit}. \quad (14)$$

Denoting R^θ as the tipping risk under an EWSys with trigger value θ, we can calculate the relative risk change as

$$\Delta_R^\theta = \frac{R^\theta - R^b}{R^b}. \quad (15)$$

This expression is positive if the EWSys increases risk and negative if it decreases risk.

2.3.2 Definition of economic value

To define a measure of the value that an EWSys provides over a situation without the EWSys, one might find it natural to define the value of the EWSys as $V^\theta - V^b$, where V^θ denotes the value of (5) under an EWSys with trigger value θ. However, that quantity is not invariant under changes to the utility function that preserve risk preferences. To get a measure that only depends on preferences but not the utility function representation, we can look at the action level $\hat{a}$ that makes the regulator indifferent between the following cases:

(i) taking action $\hat{a}$ in both periods without the risk of collapse.

(ii) the actual decision-problem we study (this can be the baseline scenario without EWSys or the scenario with EWSys).

For a given instantaneous utility function u, the intertemporal utility of the case (i) is $u(\hat{a}) + \beta u(\hat{a}) = (1 + \beta)u(\hat{a})$. Let the intertemporal utility of the scenario we study be V_1, cf. expression (5), i.e. all actions have already been chosen optimally.

The indifference condition is $(1 + \beta)u(\hat{a}) = V_1$, or $u(\hat{a}) = V_1/(1 + \beta)$. With $u(a) = (a/\bar{a})^\alpha$, we get $\hat{a} = \bar{a}[V_1/(1 + \beta)]^{1/\alpha}$.

We can calculate $\hat{a}_b$ for the baseline scenario and we can calculate $\hat{a}_\theta$ for the EWSys scenario with trigger value θ. It is easy to show that, in either case, $\hat{a}$ is invariant under cardinal transformations of the utility function, i.e. $\tilde{u} = \lambda u$. Note that we do not have to consider additive shifts of the utility function as we have set $u(0) = 0$.

In any case, a higher $\hat{a}$ means the situation is more valuable to the decision-maker. Because changes of the utility function changes the certainty-equivalent $\hat{a}$ of the baseline scenario, a suitable measure of the relative increase in economic value provided by the EWSys is the increase in the certainty-equivalent caused by the EWSys, relative to the baseline certainty-equivalent,

$$W^\theta = \frac{\hat{a}_\theta - \hat{a}_b}{\hat{a}_b}. \tag{16}$$

In the following, we shall first characterize the optimal solution of the management problem (5) for a given EWSys. Then, we will analyze, for given time preference and risk tolerance of the regulator, how the total risk of tipping depends on the trigger value of the EWSys. Finally, we shall present how the relative change in risk (15), and value (16) depend on the regulator's time and risk preferences.

3 Results

We solve the management problem (5) numerically, taking as input the socio-ecological system (1) and assuming a uniform initial prior $p_1(a_{crit}) = 1/\bar{a}$. We solve the equation system by backward recursion on the grid space, which consists of the action space (grid of 500 points, $a \in \{0, 0.004, \ldots, 2\}$), the tipping point space (grid of 500 points, $a_{crit} \in \{0, 0.004, \ldots, 2\}$), as well as the trigger value space (grid of 501 points, $\theta = \{-1, -0.996, \ldots, 1\}$) and the resulting family of signal functions $f_{a_{crit},\theta}$.

This allows us to characterize the regulator's second-period objective function after hearing an EWS, the regulator's second-period objective function after not hearing an EWS, and her first-period objective function, as well as the corresponding beliefs p_2^{EWS} and p_2^{noEWS}. The numerical results are presented in detail in the Appendix. Figure A2 shows how the optimal first period action depends on the trigger value θ. Figure A3 shows the objective function for the first-period action, given the preference parameters $\alpha = 0.4$ and $\beta = 0.8$. Figure A4 then zooms in on how the EWS probability q and expected second-period value depend on the first period action a_1. Figure A5 shows the second-period objective function and probability of not tipping for EWS and no EWS. Finally, Figure A6 shows the second period action choice a_2^* and second period step $a_2^* - a_1^*$ as a function of trigger value θ for different risk tolerance levels α, discount factors β and whether an EWS has been received. We summarize these insights below.

3.1 Characterization of the optimal solution

The *baseline* case without an EWSys results in an optimal action a^b that merely balances the expected benefits of a harvesting action with the expected cost of causing the collapse. The regulator will always choose the same level of action in the first and the second period. The reason why this behavior is indeed optimal is that the regulator does not receive any new information after the first period, apart from the fact that her action did not cause the collapse.

The situation is different when an EWSys exists. In this case, the regulator – conditional on not having crossed the tipping point a_{crit} already – will receive additional information from the scientific agency after the first action a_1. The regulator anticipates that the presence of EWSys will improve her management by taking a better-informed action in the second period. As a consequence, compared to the baseline case without an EWSys, the regulator will choose a lower action in the first period, that is $a_1^* \leq a^b$ (for the case $\alpha = 0.4$ and $\beta = 0.8$, compare the black line to the violet line in Figure A3).

The optimal second-period action a_2^* depends on whether or not the scientific agency reports an EWS. If it does, the regulator realizes that the tipping point a_{crit} is close and increasing the harvesting level will increase the risk of tipping, militating against a higher second-period action. At the same time, it does not make sense to reduce the level of harvesting because the regulator knows that the current level of harvesting, although triggered an EWS, did not cause the collapse. Consequently, after receiving an EWS the regulator will typically stay put, i.e., choose the same action in the second period as in the first period, $a_2^* = a_1^*$; exemplified by the blue lines in the right panel of Figure A6 when $\alpha \in \{0.2, 0.4, 0.6, 0.8\}$ and $\beta = 0.4$.

In contrast, if the scientific agency does not report an EWS, the regulator knows that a moderate increase in harvesting is unlikely to cause collapse. Hence, the regulator will find it optimal to increase the action in the second period, $a_2^* > a_1^*$ (see the red lines in the right panel of Figure A6).

Importantly, this means that the more cautious first-period choice is not due to the possibility of receiving bad news. Rather, the possibility of receiving good news and leaping forward at reduced risk afterwards is what drives the more cautious first-period action (see Figure A6). This is consistent with the literature on the quasi-option value and the value of flexibility (Henry, 1974; Arrow and Fisher, 1974): anticipating that one will receive more information on the tipping point location is a reason to make a more cautious decision (i.e., maintain flexibility and potentially make a bolder move in the second period).

A natural question is to ask how the combination of optimal first- and second-period actions impacts on the overall probability of tipping. Figure 4 shows the total risk of tipping and its constituent parts with an EWSys compared to the situation without an EWSys. Clearly, for a non-informative EWSys (trigger value θ below -0.6, approximately, for $\alpha = 0.8$ and $\beta = 0.4$) there is no difference to

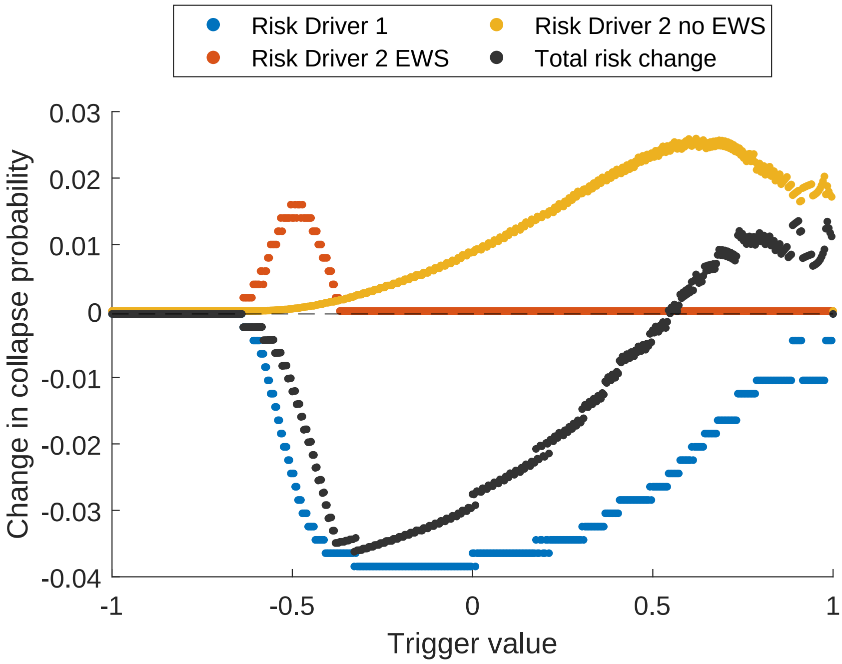


Figure 4: Disentangling the components of tipping risk. The black line shows the total change in tipping risk when comparing, for a given trigger value θ, the situation with an EWSys to the situation without an EWSys (parameter settings are risk tolerance $\alpha = 0.8$ and discount factor $\beta = 0.4$). The total change is the sum of the risk reduction due to a less risky action in period 1 (blue line), a more risky action in period 2 after receiving an EWS (red line) and after receiving no EWS (yellow line). For large values of θ, the availability of an EWSys increases the total risk of tipping.

the situation without an EWSys. As the scientific agency uses an EWSys with a higher trigger value, however, the total tipping risk decreases. After a certain point ($\theta \approx -0.4$), however, the tipping risk increases, though it is still smaller than in the situation without an EWSys. This is no longer true once θ exceeds a value of 0.5. Then, the total risk of tipping is higher with an EWSys than without an EWSys.

The three components of risk change shown in Figure 4 directly correspond to equation (14) for the change in collapse risk relative to the baseline risk. The blue line (the first component) shows how the first period action changes risk. As the optimal first action a_1^* never exceeds the baseline action a^b, this first component is non-positive. The second (red) and third (yellow) components describe how the second period actions $a_{2,EWS}^*$ and $a_{2,noEWS}^*$ change the risk relative to the situation of staying at a_1^*, respectively. Those latter components are non-negative.

We find that intermediate levels of the trigger value drive down risk the most as the first period action gets the most cautious. We also observe that the additional move despite bad news (red) can increase risk, but that this channel is limited. Intuitively, the regulator chooses a relatively low first period action to take advantage of the possibility of relatively safe increase in the second period upon hearing an EWS (bad news). But when she hears no EWS at a low trigger value, this does not mean too much (as the tigger value is low), so that she still increases the second-period action beyond the first-period action. The regulator no longer increases the second-period action beyond the first-period action once the EWSys is sufficiently informative ($\theta \approx -0.4$ for the current parameter combination). The reason for an overshooting of risk at high trigger values is the increased risk from moving after hearing good news (no EWS), and the fact that this additional risk is not counterbalanced by a more cautious first-period action.

Nevertheless, the possibility of receiving an EWS always increases the economic value that can be derived from the socio-ecological system because the EWSys provides additional information about the location of the tipping point (the regulator can always choose to ignore the information and be as well off as when no EWSys exists).

3.2 Important dimensions of EWSys design

While the above discussion was couched in terms of the regulator's objective for a given set of risk and time preferences, we now analyze how the change in tipping risk and the relative value gain from an EWSys depends on the trigger value for various combinations of risk and time preferences.

Figure 5 illustrates important dimensions of EWSys design. Each point in Figure 5a and Figure 5b shows, for a given combination of α and β, the change in risk Δ_R^θ (x-axis) and the increase in economic value W^θ (y-axis) that is resulted from the implementation of an EWSys with trigger value θ relative to the baseline case without an EWSys. Figure 5a on the left keeps the discount factor fixed at $\beta = 0.4$, and illustrates four different values of risk tolerance α (indicated by different colors). Figure 5b on the right keeps risk tolerance fixed at $\alpha = 0.4$ and

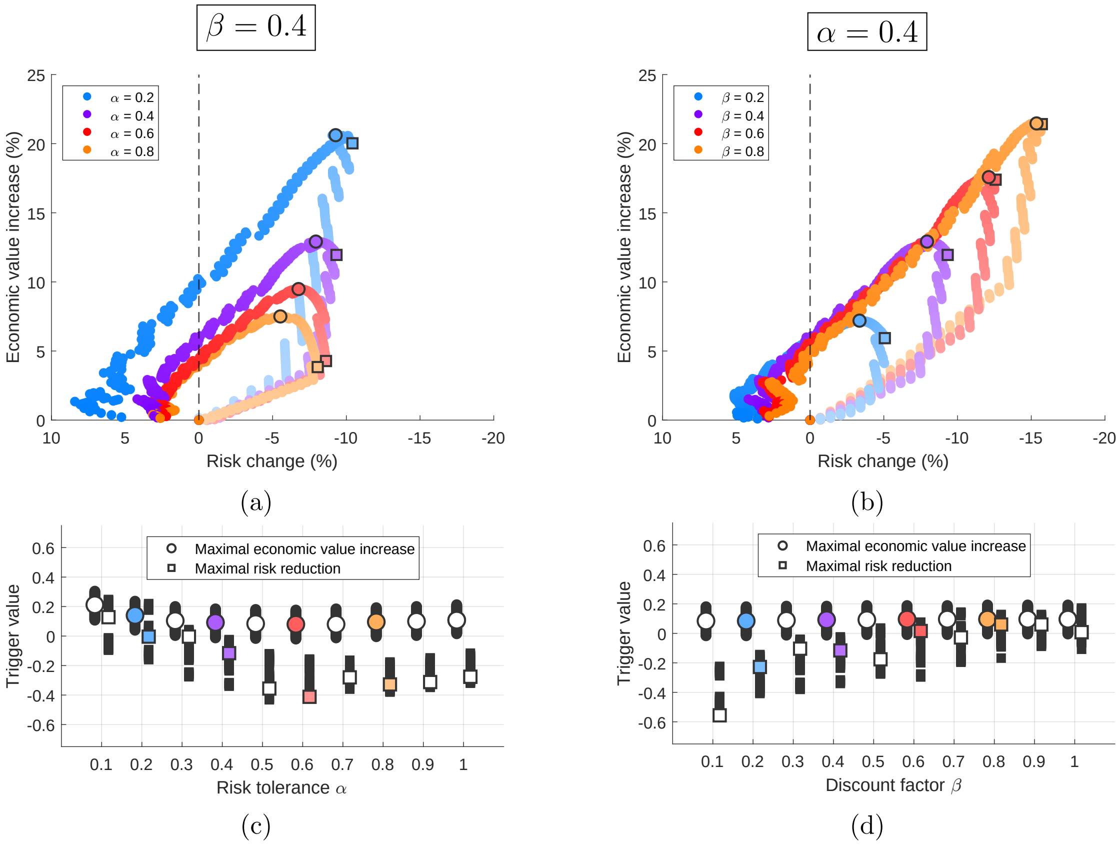


Figure 5: Relative change in risk on the horizontal axis vs. relative increase in economic value on the vertical axis for (a) fixed time preference $\beta = 0.4$ and varying risk tolerance α, and (b) varying time preference β and fixed risk tolerance $\alpha = 0.4$. Color shades indicate low trigger values (light) to high trigger values (dark). The lower panel of each figure, (c) and (d), shows the trigger values resulting in the top 10% for economic value increase (circles) and risk reduction (squares). Large symbols represent the optimal trigger value in the respective dimension and directly correspond to the symbols in the upper panel.

illustrates four different values of β. Both panels highlight the trade-off in risk and economic value as the trigger value θ changes from -1 (light shade) to 1 (dark shade), describing a counterclockwise loop.

For concreteness, let us focus on the situation where $\beta = 0.4$ and $\alpha = 0.4$ (purple points in Figure 5a). For the starting point of $\theta = -1$ (the (0,0) coordinates), there is no change relative to the situation without an EWSys because the EWSys is not informative – the tipping indicator *always* exceeds the trigger value (i.e., the

EWS is triggered too easily too early as illustrated in Figure 3b). As the trigger value θ increases, we move to the top right corner where the EWSys reduces total tipping risk and increases economic value, which is unambiguously good for management. However, there is a turning point after which a higher θ does not reduce the risk of tipping anymore but increases it (a leftward move in the risk-value coordinate system). As elaborated above, there might be cases where the total risk of tipping even exceeds the baseline risk in the absence of the EWSys (left side of the vertical dashed lines in Figure 5a and Figure 5b). Similarly, there is a turning point after which a higher θ no longer leads to a further increase in economic value but induces a downward move in the risk-economic value coordinate system.

This counterclockwise loop traced by changing the trigger value is a generic feature. A higher β (reflecting a higher relative weight placed on the future) implies that the EWSys adds more value, which is intuitive. Interestingly, a higher α (more risk tolerance) does not necessarily imply an increased total risk of tipping, as this is the result of the complex interplay of the optimal first-period action a_1^*, the learning that takes place due to that action for the given EWSys, and the optimal second-period action a_2^* (Section 2.3.1).

It is natural to mark the point at which increasing the trigger value θ no longer increases economic value (highlighted with a black outer circle in Figure 5a and Figure 5b) Similarly, we mark the point at which increasing the trigger value no longer decreases tipping risk (highlighted with a black square in Figure 5a and Figure 5b). We plot the trigger values that maximize economic value and the trigger values that minimize tipping risk for different values of risk tolerance (Figure 5c) and discount factor (Figure 5d). The figures clearly show that the best trigger values that achieve these different targets are not the same. In fact, even the trigger values that result in the top 10% value increase and the ones that result in the top 10% risk reduction do not overlap in many cases. Hence, even in this stylized model, there is a tension between the maximum increase in economic value and the maximum risk reduction.

4 Discussion

Our study develops a framework to complement the emerging knowledge on detecting early warning signals (EWS) with a better understanding of how and why early warning systems (EWSys) can improve management: the mere existence of an EWSys for detecting the tipping point invokes more cautious initial actions. If, in fact, an EWS is received, it affirms that the caution was warranted, while not receiving an EWS implies that it is unlikely that a tipping point is imminent (i.e., no news is good news). This may encourage the regulator to take a subsequent bolder action which may – in some circumstances – lead to an increased risk of tipping compared to a situation where no EWSys is available. Yet, an EWSys is always economically valuable, as the additional information can simply be disregarded.

Nevertheless, the overall societal value of an EWSys can only be assessed by its contribution towards the broader management objective. There are different views in society on the objective of ecosystem management. Those concerned with the conservation of the ecosystem may prefer an EWSys with a low trigger value to minimize the risk of collapse, while others may prefer an EWSys with a high trigger value to maximize the economic value, despite the increased tipping risk. Even if society as a whole were to subscribe to the maximization of economic value, different groups could have different preferences about risk tolerance or time discounting, leading to different valuations of a given EWSys (see Figure 5). Thus, any deliberate effort to design an EWSys carries some inevitable tensions. If the EWSys cannot be designed but must be taken as is, there might still be disagreements about whether it should be used at all. Future research may focus on analyzing EWS in interactive environments of heterogeneous actors with diverging preferences.

Our work opens two important avenues for management science. First, how do real decision makers react to EWS under impending regime shift risks? For example, Seifert et al. (2023) experimentally study how participants' under- and overreaction to dynamic regime shift risk depends on the initial prior of the decision-makers. Given our results, it is an open question whether EWS would dampen or amplify such behavioral effects. Second, how would the availability of EWSys impact decision making when the regulator and the scientific agency are strategic actors? In a recent contribution, Alizamir et al. (2020) highlight the role of the sender's credibility for its ability to induce the receiver to take certain actions. More generally, embedding the full potential to receive EWS in strategic settings will likely be a formidable challenge, but its solution promises significant contribution to both theoretical (Kamenica and Gentzkow, 2011; Bergemann and Morris, 2016) and experimental (Bolton and Katok, 2018) approaches to Bayesian persuasion.

A central feature of our model is the direct relationship between human actions and the occurrence of the tipping point. This feature characterizes several real-world problems, such as fisheries, forests, or livestock management. In other applications, however, human control is less direct. For many conservation problems species may be threatened not only by hunting and poaching, but also by habitat loss or climate change (many of these threats are indirectly caused by humans, but on different temporal and spatial scales). Better understanding of such uncontrollable processes and how they interact with human actions is a key task for future research.[7]

Finally, we have assumed that the EWS is binary and not continuous. This is in line with the existing literature on EWS (Scheffer et al., 2009; Dakos et al., 2015; Bury et al., 2021; Boettiger and Hastings, 2012) that only considers whether an EWS is sent or not. Such a setup could reflect a situation where the scientific agency has full information, but the regulator's capacity (time or money) to process it, is limited. Alternatively, the setup could reflect a data-poor environment where the scientific agency itself has only limited information or a system that is too complex and noisy to accurately monitor the tipping indicator (Chen and Tung, 2023; Boers, 2023). In principle, however, one could well think that there is more nuance to the EWS. Many practical applications, for example, work with a "traffic light system" where there is an intermediate, yellow, stage of warning in between the two extremes of warning (red), and no warning (green). A natural question

[7] Another key assumption of our model is that the tipping risk is linked only to current actions. However, in a number of systems the risk of tipping is a function of both current and past actions (Liski and Salanié, 2019; Crépin and Nævdal, 2019). In such applications, current actions have to be taken before all consequences of past actions are realized. In this case, our proposed framework can be further developed to account for the interaction of past and future tipping risk affects the value that can be derived from an EWSys, providing an interesting link to the literature on attribution (Stott et al., 2004; Otto, 2016).

for future research is how the granularity of the EWS affects the EWSys, and, in turn, its optimal design.

Despite these limitations and simplifications, our framework can pave the way for fruitful collaborations across disciplines such as ecology, economics, engineering, management and psychology. It builds a bridge to the next generation of EWS research that can now assess the risk and value of an EWS in concrete empirical applications by showing how it can be used in a management setting, and, depending on the different views in society on the objective of ecosystem management, when and why it can improve decisions in the presence of tipping points. As a result, we have made the translation from detection to operationalization readily available for use by practitioners. While our stylized socio-ecological system is generic, it can be tailored to specific real-world decision-making problems addressing sustainability and the best use of limited human and natural resources.

References

Saed Alizamir, Francis de Véricourt, and Shouqiang Wang. Warning against recurring risks: An information design approach. *Management Science*, 66(10):4612–4629, 2020.

Kenneth J. Arrow and Anthony C. Fisher. Environmental Preservation, Uncertainty, and Irreversibility. *Quarterly Journal of Economics*, 88(2):312–319, May 1974. ISSN 00335533.

Dirk Bergemann and Stephen Morris. Information design, bayesian persuasion, and bayes correlated equilibrium. *American Economic Review*, 106(5):586–91, May 2016. doi: 10.1257/aer.p20161046. URL https://www.aeaweb.org/articles?id=10.1257/aer.p20161046.

Niklas Boers. Observation-based early-warning signals for a collapse of the atlantic meridional overturning circulation. *Nature Climate Change*, 11(8):680–688, 2021.

Niklas Boers. Reply to: Evidence lacking for a pending collapse of the atlantic meridional overturning circulation. *Nature Climate Change*, 2023. ISSN 1758-6798. doi: 10.1038/s41558-023-01878-z. URL https://doi.org/10.1038/s41558-023-01878-z.

Carl Boettiger and Alan Hastings. Quantifying limits to detection of early warning for critical transitions. *Journal of The Royal Society Interface*, 9(75):2527–2539, 2012. doi: 10.1098/rsif.2012.0125. URL https://royalsocietypublishing.org/doi/abs/10.1098/rsif.2012.0125.

Gary E. Bolton and Elena Katok. Cry wolf or equivocate? credible forecast guidance in a cost-loss game. *Management Science*, 64(3):1440–1457, 2018. doi: 10.1287/mnsc.2016.2645. URL https://doi.org/10.1287/mnsc.2016.2645.

Chris A. Boulton, Timothy M. Lenton, and Niklas Boers. Pronounced loss of amazon rainforest resilience since the early 2000s. *Nature Climate Change*, 12(3):271–278, 2022. ISSN 1758-6798. doi: 10.1038/s41558-022-01287-8. URL https://doi.org/10.1038/s41558-022-01287-8.

O. Bruzzone and M.H. Easdale. Rhythm of change of trend-cycles of vegetation dynamics as an early warning indicator for land management. *Ecological Indicators*, 126:107663, 2021. ISSN 1470-160X. doi: https://doi.org/10.1016/j.ecolind.2021.107663. URL https://www.sciencedirect.com/science/article/pii/S1470160X21003289.

T. M. Bury, C. T. Bauch, and M. Anand. Detecting and distinguishing tipping points using spectral early warning signals. *Journal of The Royal Society Interface*, 17(170):20200482, 2020. doi: 10.1098/rsif.2020.0482. URL https://royalsocietypublishing.org/doi/abs/10.1098/rsif.2020.0482.

Thomas M. Bury, R. I. Sujith, Induja Pavithran, Marten Scheffer, Timothy M. Lenton, Madhur Anand, and Chris T. Bauch. Deep learning for early warning signals of tipping points. *Proceedings of the National Academy of Sciences*, 118(39), 2021. ISSN 0027-8424. doi: 10.1073/pnas.2106140118. URL https://www.pnas.org/content/118/39/e2106140118.

Yongyang Cai and Thomas S. Lontzek. The social cost of carbon with economic and climate risks. *Journal of Political Economy*, 127(6):2684–2734, 2019. doi: 10.1086/701890. URL https://doi.org/10.1086/701890.

Yongyang Cai, Kenneth L. Judd, Timothy M. Lenton, Thomas S. Lontzek, and Daiju Narita. Environmental tipping points significantly affect the cost-benefit assessment of climate policies. *Proceedings of the National Academy of Sciences*, 112(15):4606–4611, 2015. ISSN 0027-8424. doi: 10.1073/pnas.1503890112. URL `https://www.pnas.org/content/112/15/4606`.

S. R. Carpenter, J. J. Cole, M. L. Pace, R. Batt, W. A. Brock, T. Cline, J. Coloso, J. R. Hodgson, J. F. Kitchell, D. A. Seekell, L. Smith, and B. Weidel. Early warnings of regime shifts: A whole-ecosystem experiment. *Science*, 332(6033):1079–1082, 2011. ISSN 0036-8075. doi: 10.1126/science.1203672. URL `https://science.sciencemag.org/content/332/6033/1079`.

Xianyao Chen and Ka-Kit Tung. Evidence lacking for a pending collapse of the atlantic meridional overturning circulation. *Nature Climate Change*, 2023. ISSN 1758-6798. doi: 10.1038/s41558-023-01877-0. URL `https://doi.org/10.1038/s41558-023-01877-0`.

Christopher F. Clements, John M. Drake, Jason I. Griffiths, Arpat Ozgul, Associate Editor: Egbert H. van Nes, and Editor: Susan Kalisz. Factors influencing the detectability of early warning signals of population collapse. *The American Naturalist*, 186(1):50–58, 2015. ISSN 00030147, 15375323. URL `http://www.jstor.org/stable/10.1086/681573`.

Christopher F Clements, Julia L Blanchard, Kirsty L Nash, Mark A Hindell, and Arpat Ozgul. Body size shifts and early warning signals precede the historic collapse of whale stocks. *Nature ecology & evolution*, 1(7):1–6, 2017.

Anne-Sophie Crépin and Eric Nævdal. Inertia risk: Improving economic models of catastrophes. *The Scandinavian Journal of Economics*, n/a(n/a):1–27, 2019. doi: 10.1111/sjoe.12381. URL `https://www.onlinelibrary.wiley.com/doi/abs/10.1111/sjoe.12381`.

V. Dakos, C. A. Boulton, J. E. Buxton, J. F. Abrams, D. I. Armstrong McKay, S. Bathiany, L. Blaschke, N. Boers, D. Dylewsky, C. López-Martínez, I. Parry, P. Ritchie, B. van der Bolt, L. van der Laan, E. Weinans, and S. Kéfi. Tipping point detection and early-warnings in climate, ecological, and human systems. *EGUsphere*, 2023:1–35, 2023. doi: 10.5194/egusphere-2023-1773. URL `https://egusphere.copernicus.org/preprints/2023/egusphere-2023-1773/`.

Vasilis Dakos, Marten Scheffer, Egbert H. van Nes, Victor Brovkin, Vladimir Petoukhov, and Hermann Held. Slowing down as an early warning signal for abrupt climate change. *Proceedings of the National Academy of Sciences*, 105(38):14308–14312, 2008. ISSN 0027-8424. doi: 10.1073/pnas.0802430105. URL `https://www.pnas.org/content/105/38/14308`.

Vasilis Dakos, Stephen R. Carpenter, William A. Brock, Aaron M. Ellison, Vishwesha Guttal, Anthony R. Ives, Sonia Kéfi, Valerie Livina, David A. Seekell, Egbert H. van Nes, and Marten Scheffer. Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data. *PLOS ONE*, 7(7):e41010, July 2012. doi: 10.1371/journal.pone.0041010. URL `https://doi.org/10.1371/journal.pone.0041010`.

Vasilis Dakos, Stephen R. Carpenter, Egbert H. van Nes, and Marten Scheffer. Resilience indicators: prospects and limitations for early warnings of regime shifts. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 370(1659), 2015. ISSN 0962-8436. doi: 10.1098/rstb.2013.0263.

Heiman Dianat, Suzanne Wilkinson, Peter Williams, and Hamed Khatibi. Choosing a holistic urban resilience assessment tool. *International Journal of Disaster Risk Reduction*, 71:102789, 2022. ISSN 2212-4209. doi: https://doi.org/10.1016/j.ijdrr.2022.102789. URL https://www.sciencedirect.com/science/article/pii/S2212420922000085.

Florian K. Diekert. Threatening thresholds? the effect of disastrous regime shifts on the non-cooperative use of environmental goods and services. *Journal of Public Economics*, 147:30 – 49, 2017. ISSN 0047-2727. doi: http://dx.doi.org/10.1016/j.jpubeco.2017.01.004. URL http://www.sciencedirect.com/science/article/pii/S0047272717300130.

Simon Dietz, James Rising, Thomas Stoerk, and Gernot Wagner. Economic impacts of tipping points in the climate system. *Proceedings of the National Academy of Sciences*, 118(34): e2103081118, 2021. doi: 10.1073/pnas.2103081118. URL https://www.pnas.org/doi/abs/10.1073/pnas.2103081118.

Peter Ditlevsen and Susanne Ditlevsen. Warning of a forthcoming collapse of the atlantic meridional overturning circulation. *Nature Communications*, 14(1):4254, 2023. ISSN 2041-1723. doi: 10.1038/s41467-023-39810-w. URL https://doi.org/10.1038/s41467-023-39810-w.

Yuval Hart, Maryam Vaziri-Pashkam, and L. Mahadevan. Early warning signals in motion inference. *PLOS Computational Biology*, 16(5):1–16, 05 2020. doi: 10.1371/journal.pcbi.1007821. URL https://doi.org/10.1371/journal.pcbi.1007821.

Hermann Held and Thomas Kleinen. Detection of climate system bifurcations by degenerate fingerprinting. *Geophysical Research Letters*, 31(23), 2004. doi: 10.1029/2004GL020972. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2004GL020972.

Marieke A. Helmich, Arnout C. Smit, Laura F. Bringmann, Marieke J. Schreuder, Albertine J. Oldehinkel, Marieke Wichers, and Evelien Snippe. Detecting impending symptom transitions using early-warning signals in individuals receiving treatment for depression. *Clinical Psychological Science*, 0(0):21677026221137006, 2022. doi: 10.1177/21677026221137006. URL https://doi.org/10.1177/21677026221137006.

Claude Henry. Investment Decisions Under Uncertainty: The 'Irreversibility Effect.'. *American Economic Review*, 64(6):1006–1012, December 1974. ISSN 00028282.

Marina Hirota, Milena Holmgren, Egbert H. Van Nes, and Marten Scheffer. Global resilience of tropical forest and savanna to critical transitions. *Science*, 334(6053):232–235, 2011. doi: 10.1126/science.1210657. URL https://www.science.org/doi/abs/10.1126/science.1210657.

Emir Kamenica and Matthew Gentzkow. Bayesian persuasion. *American Economic Review*, 101 (6):2590–2615, October 2011. doi: 10.1257/aer.101.6.2590. URL https://www.aeaweb.org/articles?id=10.1257/aer.101.6.2590.

D Lemoine and C Traeger. Watch your step: Optimal policy in a tipping climate. *American Economic Journal: Economic Policy,*, 6(1):137–166, 2014.

Timothy M Lenton, Johan Rockström, Owen Gaffney, Stefan Rahmstorf, Katherine Richardson, Will Steffen, and Hans Joachim Schellnhuber. Climate tipping points – too risky to bet against. *Nature*, 575:592–595, 2019.

Matti Liski and Salanìè. Tipping points, delays, and the control of catastrophestipping points, delays, and the control of catastrophes. unpublished manuscript, 2019.

Yanlan Liu, Mukesh Kumar, Gabriel G Katul, and Amilcare Porporato. Reduced resilience as an early warning signal of forest mortality. *Nature Climate Change*, 9(11):880–885, 2019.

Thomas E. Lovejoy and Carlos Nobre. Amazon tipping point. *Science Advances*, 4(2), 2018. doi: 10.1126/sciadv.aat2340. URL https://advances.sciencemag.org/content/4/2/eaat2340.

David I. Armstrong McKay, Arie Staal, Jesse F. Abrams, Ricarda Winkelmann, Boris Sakschewski, Sina Loriani, Ingo Fetzer, Sarah E. Cornell, Johan Rockström, and Timothy M. Lenton. Exceeding 1.5c global warming could trigger multiple climate tipping points. *Science*, 377(6611):eabn7950, 2022. doi: 10.1126/science.abn7950. URL https://www.science.org/doi/abs/10.1126/science.abn7950.

Eric Nævdal. Dynamic optimisation in the presence of threshold effects when the location of the threshold is uncertain – with an application to a possible disintegration of the western antarctic ice sheet. *Journal of Economic Dynamics and Control*, 30(7):1131–1158, July 2006. ISSN 0165-1889. URL http://www.sciencedirect.com/science/article/pii/S0165188905000928.

Duncan A. O'Brien and Christopher F. Clements. Early warning signal reliability varies with covid-19 waves. *Biology Letters*, 17(12):20210487, 2021. doi: 10.1098/rsbl.2021.0487. URL https://royalsocietypublishing.org/doi/abs/10.1098/rsbl.2021.0487.

Friederike EL Otto. The art of attribution. *Nature Climate Change*, 6(4):342–343, 2016.

Fotios Petropoulos, Daniele Apiletti, Vassilios Assimakopoulos, Mohamed Zied Babai, Devon K. Barrow, Souhaib Ben Taieb, Christoph Bergmeir, Ricardo J. Bessa, Jakub Bijak, John E. Boylan, Jethro Browell, Claudio Carnevale, Jennifer L. Castle, Pasquale Cirillo, Michael P. Clements, Clara Cordeiro, Fernando Luiz Cyrino Oliveira, Shari De Baets, Alexander Dokumentov, Joanne Ellison, Piotr Fiszeder, Philip Hans Franses, David T. Frazier, Michael Gilliland, M. Sinan Gönül, Paul Goodwin, Luigi Grossi, Yael Grushka-Cockayne, Mariangela Guidolin, Massimo Guidolin, Ulrich Gunter, Xiaojia Guo, Renato Guseo, Nigel Harvey, David F. Hendry, Ross Hollyman, Tim Januschowski, Jooyoung Jeon, Victor Richmond R. Jose, Yanfei Kang, Anne B. Koehler, Stephan Kolassa, Nikolaos Kourentzes, Sonia Leva, Feng Li, Konstantia Litsiou, Spyros Makridakis, Gael M. Martin, Andrew B. Martinez, Sheik Meeran, Theodore Modis, Konstantinos Nikolopoulos, Dilek Önkal, Alessia Paccagnini, Anastasios Panagiotelis, Ioannis Panapakidis, Jose M. Pavía, Manuela Pedio, Diego J. Pedregal, Pierre Pinson, Patrícia Ramos, David E. Rapach, J. James Reade, Bahman Rostami-Tabar, Michał Rubaszek, Georgios Sermpinis, Han Lin Shang, Evangelos Spiliotis, Aris A. Syntetos, Priyanga Dilini Talagala, Thiyanga S. Talagala, Len Tashman, Dimitrios Thomakos, Thordis Thorarinsdottir, Ezio Todini, Juan Ramón Trapero Arenas, Xiaoqian Wang, Robert L. Winkler, Alisa Yusupova, and Florian Ziel. Forecasting: theory and practice. *International Journal of Forecasting*, 38(3):705–871, 2022. ISSN 0169-2070. doi: https://doi.org/10.1016/j.ijforecast.2021.11.001. URL https://www.sciencedirect.com/science/article/pii/S0169207021001758.

Stephen Polasky, Aart de Zeeuw, and Florian Wagener. Optimal management with potential regime shifts. *Journal of Environmental Economics and Management*, 62(2):229 – 240, 2011. ISSN 0095-0696. doi: DOI:10.1016/j.jeem.2010.09.004. URL http://www.sciencedirect.com/science/article/pii/S0095069611000556.

Sumit Sarkar and Ram S Sriram. Bayesian models for early warning of bank failures. *Management Science*, 47(11):1457–1475, 2001.

Marten Scheffer, Jordi Bascompte, William A. Brock, Victor Brovkin, Stephen R. Carpenter, Vasilis Dakos, Hermann Held, Egbert H. van Nes, Max Rietkerk, and George Sugihara. Early-warning signals for critical transitions. *Nature*, 461(7260):53–59, September 2009. ISSN 0028-0836. URL http://dx.doi.org/10.1038/nature08227.

Matthias Seifert, Canan Ulu, and Sreyaa Guha. Decision making under impending regime shifts. *Management Science*, 69(10):6165–6180, 2023. doi: 10.1287/mnsc.2022.4661. URL https://doi.org/10.1287/mnsc.2022.4661.

Peter A Stott, Dáithí A Stone, and Myles R Allen. Human contribution to the european heatwave of 2003. *Nature*, 432(7017):610–614, 2004.

Yacov Tsur and Amos Zemel. Uncertainty and irreversibility in groundwater resource management. *Journal of Environmental Economics and Management*, 29(2):149 – 161, 1995. ISSN 0095-0696. doi: DOI:10.1006/jeem.1995.1037. URL http://www.sciencedirect.com/science/article/pii/S0095069685710376.

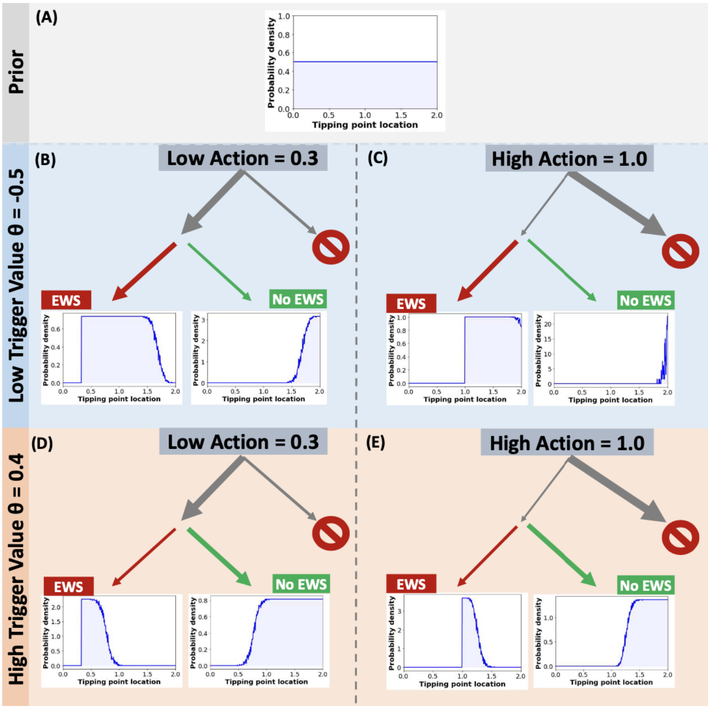
Ingrid A. van de Leemput, Marieke Wichers, Angélique O. J. Cramer, Denny Borsboom, Francis Tuerlinckx, Peter Kuppens, Egbert H. van Nes, Wolfgang Viechtbauer, Erik J. Giltay, Steven H. Aggen, Catherine Derom, Nele Jacobs, Kenneth S. Kendler, Han L. J. van der Maas, Michael C. Neale, Frenk Peeters, Evert Thiery, Peter Zachar, and Marten Scheffer. Critical slowing down as early warning for the onset and termination of depression. *Proceedings of the National Academy of Sciences*, 111(1):87–92, 2014. doi: 10.1073/pnas.1312114110. URL https://www.pnas.org/doi/abs/10.1073/pnas.1312114110.

Rudi Voss and Martin Quaas. Fisheries management and tipping points: Seeking optimal management of eastern baltic cod under conditions of uncertainty about the future productivity regime. *Natural Resource Modeling*, 35:e12336, 2022.

Haoyu Wen, Massimo Pica Ciamarra, and Siew Ann Cheong. How one might miss early warning signals of critical transitions in time series data: A systematic study of two major currency pairs. *PLOS ONE*, 13(3):1–22, 03 2018. doi: 10.1371/journal.pone.0191439. URL https://doi.org/10.1371/journal.pone.0191439.

Appendix 1 Illustration of belief updating

Here we illustrate how the regulator updates her belief about the location of the tipping point. In our stylized model, the location of the tipping point corresponds to the highest action that the system can sustain. We illustrate four cases: the regulator chooses either a high or a low action, and the EWSys has either a high or a low trigger value. The starting point for all four cases is the same: the regulator holds a uniform prior p_1 about the location of the tipping point (with $\int_0^{\bar{a}} p_1(a_{crit})da_{crit} = 1$, where $\bar{a} = 2$ is the maximum harvesting action), Panel (A).



Consider first the two cases where the trigger value of the EWSys is low ($\theta = -0.5$), Panels (B) and (C). Choosing a low action (a=0.3) implies a small probability of causing the collapse of the system whereas choosing a high action (a=1.0) implies a large probability of causing collapse (shown by the thickness of the grey arrows). Conditional on not causing the collapse, the regulator learns that the respective action level is safe. Consequently, the probability that the tipping point is at one of the remaining potential locations must be higher. When the regulator receives an EWS, the probability that the tipping point is close is high, relative to the probability that the tipping point is far. On the other hand, if the regulator receives no EWS, the probability that the tipping point is close is low: no news is good news.

The same logic holds when the trigger value of the EWSys is high (θ=0.4), Panels (D) and (E). While the probability of collapse conditional on action is the same, the probability to receive an EWS and the resulting posteriors differ.

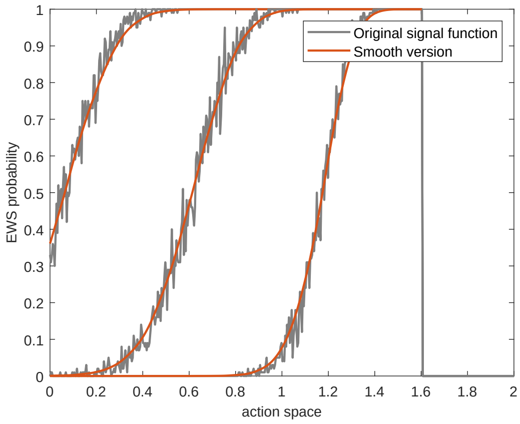
We now compare how a regulator choosing a low action updates her belief upon *not* receiving an EWS from an EWSys with low or high trigger value, Panel (B) vs. Panel (D). Recall that not receiving an EWS is good news (indicated by the green arrow). In the case of a low trigger value, the probability of not receiving an EWS is low, but the regulator becomes optimistic (in terms of assigning a lower probability that a given action causes collapse) for a relatively large range. In contrast, with a high trigger value, it is likely that the regulator receives no EWS. Correspondingly, when the regulator does not receive an EWS, she becomes optimistic only for a small range in front of the current position.

Appendix 2 Numerical implementation

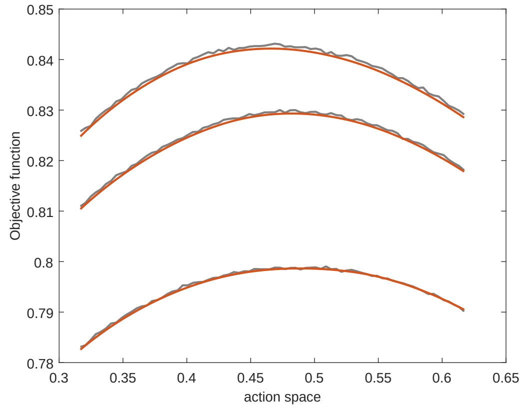
We numerically solve the equation system (1), (3), and (5), by backward recursion on the grid space. The grid space consists of the action space (grid of 500 points, $a \in \{0, 0.004, \ldots, 2\}$), the tipping point space (grid of 500 points, $a_{crit} \in \{0, 0.004, \ldots, 2\}$), as well as the trigger value space (grid of 501 points, $\theta = \{-1, -0.996, \ldots, 1\}$) and the resulting family of signal functions $f_{a_{crit},\theta}$.

That is, for each feasible point, we calculate the value of regulator's second-period objective function after hearing an EWS, and the value of the regulator's second-period objective function after not hearing an EWS, as well as the corresponding beliefs p_2^{noEWS} and p_2^{noEWS}. We then plug these values into the regulator's first-period maximization problem to find the combination of first-period action a_1^*, second-period action under no news $a_{2,noEWS}^*$, and second-period action with EWS $a_{2,EWS}^*$ that maximizes (5).

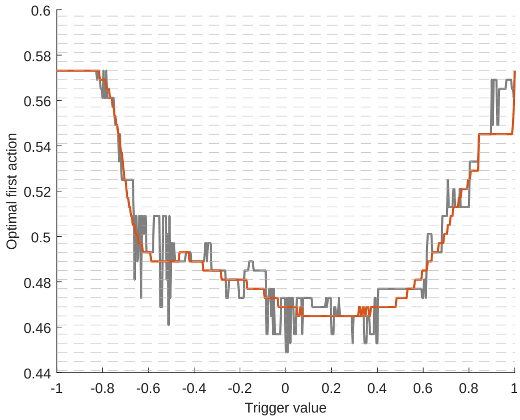
As the calculated EWS probability that results from a given combination of action, tipping point, and trigger value is the outcome of a specific realisation of noise terms, we apply moving average to the data to work with smooth versions of the signal function, see footnote 5 and Figure A1. This procedure reduces noise in the outcome variables, see panels (c) and (d), it has no discernible impact on the firs-period objective function (see panel (b) of Figure A1).



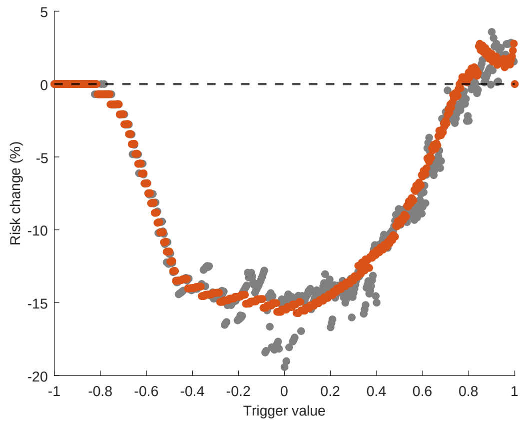
(a) Signal functions for fixed tipping point $a_{crit} = 1.6032$ and trigger value $\theta = -0.56$ (left), -0.26 (middle), and 0.16 (right).



(b) First period objective functions for trigger value $\theta = -0.56$ (bottom), -0.26 (middle), and 0.16 (top).



(c) Optimal first period action as a function of trigger value θ. The horizontal dashed lines show the action space grid.

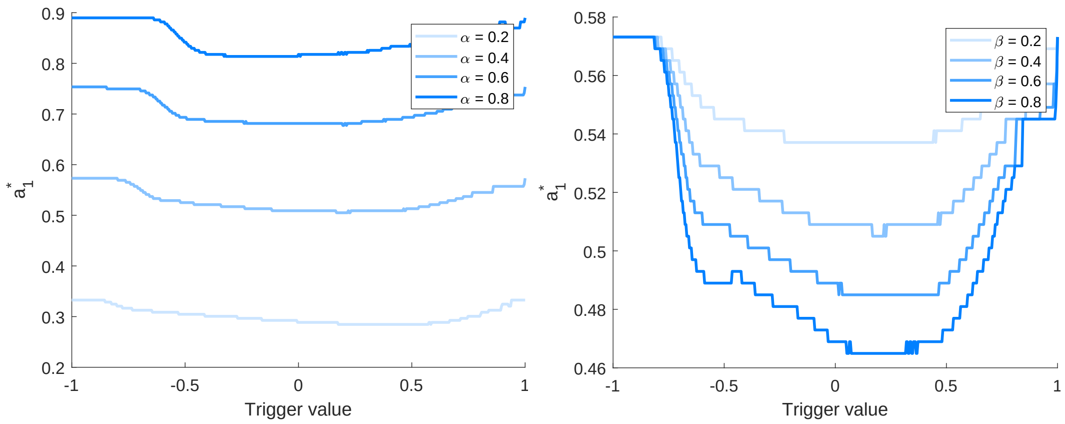


(d) Relative risk change implied by the optimal actions as a function of trigger value θ.

Figure A1: Smoothing the signal functions. Comparison of the original signal functions (in gray) with the smooth signal functions (in red). The smooth version is a moving average of the original signal functions. The preference parameters in panels (c) and (d) are $\alpha = 0.4$ and $\beta = 0.8$.

Appendix 3 Detailed analysis

Figure A2 shows the optimal first-period action a_1^* with an EWSys as a function of its trigger value θ. First, both very low and high trigger values are non-informative and hence equivalent to the absence of an EWSys. Therefore the optimal actions with such an EWSys are equivalent to the baseline optimal action a^b. This baseline level is independent of the discount factor β (right panel) but strongly dependent on risk tolerance α (left panel), see equation (11). Second, intermediate trigger values lead to decreases in the optimal first period action in order to wait for the information provided by the EWSys. This wait-and-see effect is more pronounced when the future matters more (high discount factor, right panel).



(a) Discount factor $\beta = 0.4$.

(b) Risk tolerance $\alpha = 0.4$.

Figure A2: Optimal first-period action a_1^* as a function of trigger value θ for different levels of risk tolerance (left panel) and different discount factors (right panel). Note the different scale of the vertical axis.

Figure A3 shows the objective function for the first-period action, given the preference parameters $\alpha = 0.4$ and $\beta = 0.8$.

The objective function combines expected second-period value, first-period utility, and probability of survival, see equation (5). The blue line (red line) represents the objective function for the hypothetical first-period choice when the decision-maker knows for sure that she will receive bad news (good news). The blue line peaks where the black line peaks. This indicates that what causes the more cautious first-period action with an EWSys is not the possibility to receive bad news;

rather, the possibility to observe good news and then leap forward drives the more cautious choice. This is why the red line peaks for a low first-period action a_1. For the selected trigger values (which are fairly small), an increase in trigger value tends to reduce the optimal first-period action. This pattern would be reversed for larger trigger values, cf. Figure A2.

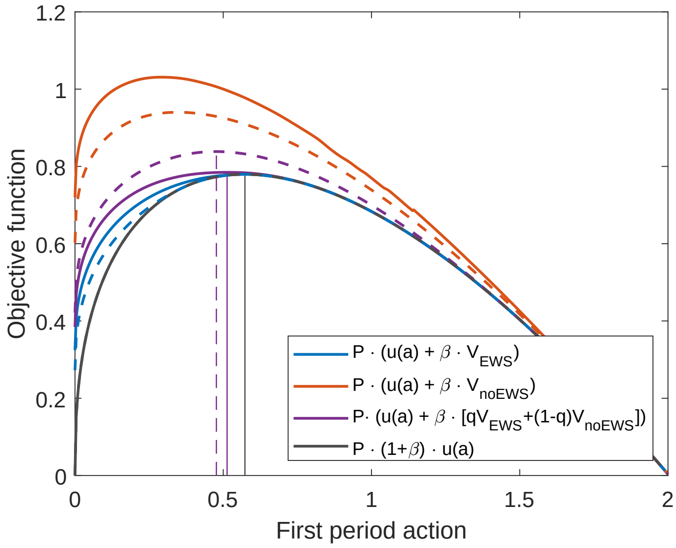
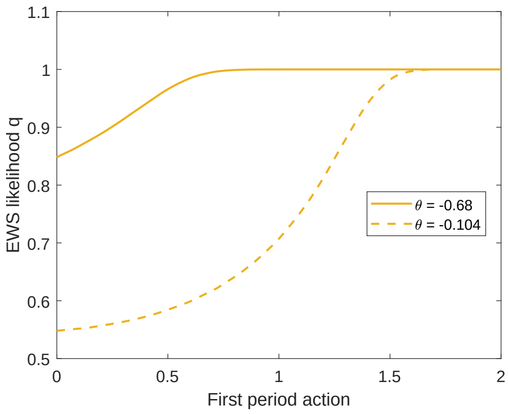


Figure A3: Objective function for choosing the optimal first-period action (see (5)). Trigger values are $\theta = -0.68$ (solid lines) and $\theta = -0.1040$ (dashed lines), and preference parameters are $\alpha = 0.4$ and $\beta = 0.8$.

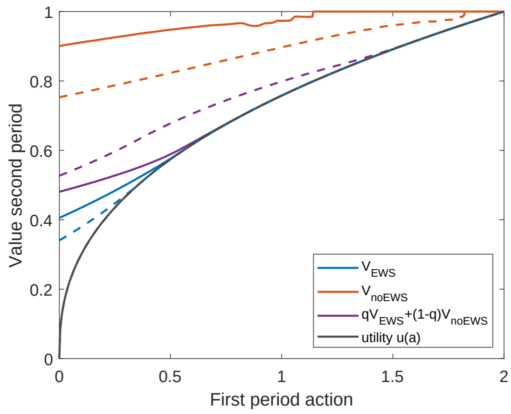
Figure A4 zooms in on two subcomponents of the objective function. Panel (a) shows the EWS probability q and panel (b) shows the expected second-period value as functions of a_1.

In panel (a) we see that the likelihood to receive an EWS increases with the first period action. The lower the trigger value, the more sensitive the EWSys and therefore the higher the likelihood to receive an EWS.

In panel (b) we see that the expected value coincides with the utility function for large first-period actions. The reason is that receiving the EWS is very likely (see panel (a)) and the best action $a^*_{2,EWS}$ after bad news is (typically) to stay, so that expected value of first-period action a_1 coincides with $u(a_1)$. That the expected value (purple line) never falls below the utility function (black line) indicates that the information provided by the EWSys is always (weakly) valuable.



(a) EWS probability q.



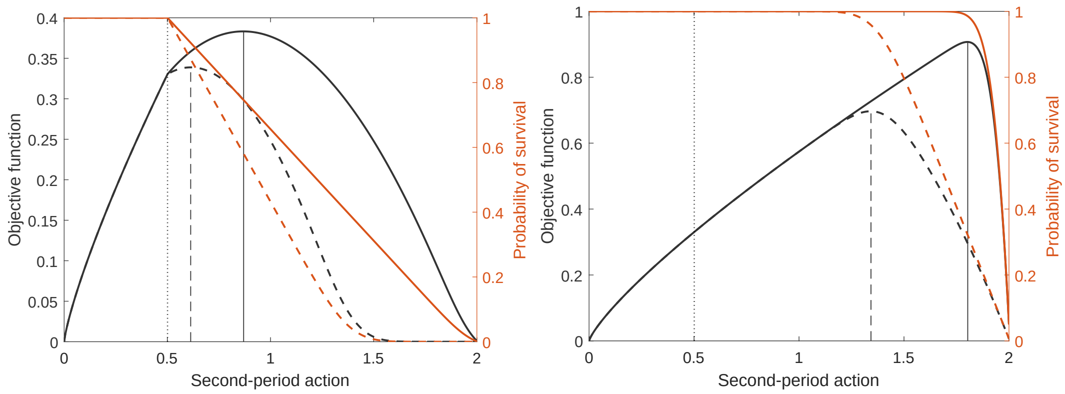
(b) Expected second-period value.

Figure A4: Panel (a): EWS probability q, cf. expression (8), as a function of the first period action, conditional on survival. Panel (b): Second-period values of the objective function conditional on survival, cf. (5), and its expected value (purple line). For comparison we show the utility function that represents (conditional on survival) the second period value without an EWSys (black line). In both panels, we have $\alpha = 0.4$ and $\beta = 0.8$.

Figure A5 shows the objective functions for choosing the optimal second period actions, conditional on receiving an EWS or not receiving an EWS after an arbitrary first period action.

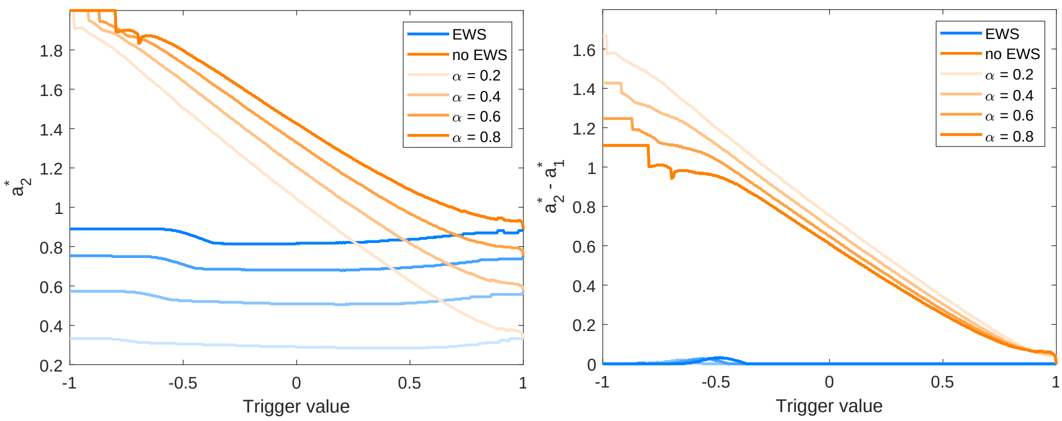
The objective function (in black, left axis) that determines the optimal second-period action (indicated by a vertical line) is the product of survival probability (decreasing in a_2) and utility function (increasing in a_2), see (5). We plot these functions for two different trigger values, $\theta = -0.680$ (solid lines) and $\theta = -0.104$ (dashed lines). We also show the probability of survival $P(a_2)$ as a function of the second-period action a_2 (in red, right axis). Survival is guaranteed if the second-period action does not exceed the first-period action.

The higher trigger value (dashed lines) corresponds to a less sensitive EWSys; bad news is therefore taken seriously and the optimal second-period action is more cautious than for a more sensitive EWSys (solid line). Similarly, if good news (no EWS) materializes, then this is big news if the EWSys is sensitive and therefore induces a bold move forward (solid line). We spell out in more detail how the section period action/step depends on the trigger value in Figure A6 below.



(a) After bad news ("EWS"). (b) After good news ("noEWS").

Figure A5: Objective functions for choosing the optimal second-period actions after bad news (left panel) and good news (right panel). The arbitrary current position inherited from the first period (not chosen optimally) is $a_1 = 0.501$ (indicated by the dotted vertical line). Risk tolerance is $\alpha = 0.8$. The trigger value θ is -0.68 (solid lines) and -0.104 (dashed lines).



(a) Second period choice a_2^*. (b) Second period step $a_2^* - a_1^*$.

Figure A6: Optimal second-period action after bad news (blue lines) and good news (red lines) as a function of trigger value θ for different levels of risk tolerance α (light colors correspond to lower values of α). The left panel shows absolute optimal second period actions $a^*_{2,EWS}$ and $a^*_{2,noEWS}$. The right panel shows the additional step between first and second period $a^*_{2,EWS} - a^*_1$ and $a^*_{2,noEWS} - a^*_1$. The discount factor (relevant for the choice of a^*_1) is $\beta = 0.4$.

Figure A6 shows the optimal second-period actions after bad news (blue lines) and good news (red lines) as a function of trigger value θ for different levels of risk tolerance α (light colors correspond to lower values of α). Several observations stand out: First, the usual response to bad news is to stay put (blue lines, right panel). Second, the more sensitive the EWSys (lower trigger values), the bolder the second-period action after good news. Third, higher risk tolerance implies higher actions and therefore higher risk (left panel). Fourth, lower risk tolerance implies more cautious actions in the first period, followed by a stronger adjustment in case of good news (right panel). The interaction of these effects causes the non-monotone pattern in tipping risk (see Figure 4).