Original Research Article

The inevitability of surprise in agroecosystems

John Vandermeer

Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI 48109, USA

ARTICLE INFO

Article history:
Received 22 May 2011
Received in revised form 6 October 2011
Accepted 7 October 2011
Available online 2 November 2011

Keywords:
Complex systems
Critical transformation
Chaos

ABSTRACT

Many critical transformations of ecosystems contain advanced signals of their imminence, but it is also true that many critical transformations can be shown to contain no such signal, at least with the sorts of data normally available to field workers. This paper explores some generalized theoretical structures and distinguishes between those that may provide a signal that could be used to predict a critical transformation and those that, by their very nature, do not provide such a clue. I argue that it is almost certain that such completely unpredictable structures will sometimes emerge from systems that tend to be as complicated as agroecosystems, in both their natural and social sciences manifestations. Precaution should thus be taken when designing agroecosystems.

© 2011 Published by Elsevier B.V.

1. Introduction

It has long been familiar knowledge that ecosystems are complex and, partially due to this complexity, frequently undergo rapid and seemingly unexpected changes. More recently we have discovered a variety of theoretical structures that seem to illustrate detailed, yet very general, mechanisms whereby such changes may occur (see reviews in Scheffer, 2009; Scheffer et al., 2009). Furthermore, these theoretical structures frequently provide clues that may enable ecosystem managers to predict when a major shift in ecosystem properties is on the horizon, a “regime shift” or “critical transition.” While most of this literature focuses on non-agricultural ecosystems, its relevance for agroecosystems can hardly be exaggerated. Pest managers routinely search for “action thresholds” to plan control activities. Small-scale farmers look to crop signals to indicate a piece of land should be left fallow next season. Although such searching for signals of impending ecosystem transformation is a regular activity of agroecosystem management, the recent tools developed by ecologists (Folke et al., 2007; Van Nes and Scheffer, 2005; Carpenter and Brock, 2006; Dakos et al., 2008; Guttal and Jayaprakash, 2008; Scheffer, 2009; Rietkerk et al., 2004) to more systematically recognize what may sometimes be very subtle clues, is a welcome incursion of the field of ecology into ecosystem management, especially in production ecosystems like agriculture.

Even as the search for predictive patterns is an important research activity, there is another issue that emerges from this body of work in theoretical ecology. It is true that many critical transformations contain advanced signals of their imminence, but it is also true that many critical transformations can be shown to contain no such signal. That is, by their very nature they are inherently unpredictable. Such structures inevitably generate surprise – a new pest, a collapsing market, desertified pastures, and many other possible examples come to mind. In this paper I explore some generalized theoretical structures and distinguish between those that may very well provide a signal that can be used to predict a critical transformation and those that, by their very nature, do not provide such a clue. I argue that it is almost certain that such completely unpredictable structures will sometimes emerge from systems that tend to be as complicated as agroecosystems, in both their natural and social sciences manifestations.

The issue is parallel to the problem identified in the early 1900s by John Maynard Keynes (Skidelsky, 2009). Economic systems are exceedingly complicated, and it was something of a triumph when economists of the early part of the last century were able to quantitatively assess risk. While there was no way of avoiding negative economic twists and turns, one could nevertheless put a probability on various negative outcomes and in that way quantify the risk of almost any economic activity, an attitude that can lead to the construction of ever more complicated financial instruments. Keynes noted that these new economists failed to see the critical difference between risk and uncertainty. When something totally new happens, something that had not been part of the risk assessment framework, the response is typically surprise. What Keynes noted was not that the evaluation of risk was pointless, but that surprises would always be on the horizon – that there was an inevitability of surprise.

There are sufficient theoretical structures informing ecology and economics today so as to enable some formal statements about critical transitions and possible system behaviors that could
provide advance warning that the critical transition is soon due (Scheffer, 2009; Scheffer et al., 2009). But those same formalities also enable some formal statements about surprise. Consider the two time series in Fig. 1. The first (Fig. 1a) clearly exhibits a signal that would permit one to predict that a population that had been slowly increasing the amplitude of its oscillations, suddenly becomes very erratic, signaling the impending collapse of the population. The second (Fig. 1b) has no such signal, suggesting that there is, in principle, no way to anticipate the same regime shift (i.e., population extinction) (McCann, 1994). It, of course, makes sense to search for signals of impending critical transitions in the case of the first example, but not in the second. I argue here that there are several qualitative situations, well-known in ecology, that are effectively examples of the second (Fig. 1b). These are cases in which a critical transition will occur, but in principle, that transition will not generate a prior signal. The shift will be a surprise and, under certain well-known conditions in ecology, that surprise will be inevitable.

2. Example from classic competition theory

Classical competition theory traditionally informs the issue of weed control, where management seeks to promote the interspecific competitive advantage of those weed species that exert the minimal negative effect on the crop. Assuming that the weeds arrive as a continual small seed rain on the field, the modified Lotka Volterra competition equations model the competitive interaction between two weed species, as follows

\[
\frac{dN_1}{dt} = N_1 \frac{K_1 - N_1 - \alpha_{12}N_2}{K_1}
\]

\[
\frac{dN_2}{dt} = N_2 \frac{K_2 - N_2 - \alpha_{21}N_1}{K_2},
\]

where \(N_i\) is the biomass of species \(i\), \(K_i\) is the carrying capacity of species \(i\), \(\alpha_{ij}\) is the competitive effect of species \(j\) on species \(i\). If we allow the competition coefficients to be equal and vary one of the carrying capacities \(K_i\), we can mimic the effects of some environmental change causing a slow secular change in the relative carrying capacities of the two weed species. The expectation may very well be that as the environmental factor changes, it affects the two weed species differently, and if it positively affects the more noxious weed, it is a factor of concern to the farmer. Suppose, for example that species 1 is a benign weed and species 2 is a noxious one and that competition coefficients are relatively small (< 1.0). Each year the transient behavior tends to an equilibrium point, defined by those competition coefficients. If we now increase the carrying capacity of the noxious species, the equilibrium point (which is always stable) will gradually change so that the noxious species will become more and more common. If we begin with \(K_1\) (carrying capacity of the benign species) much larger than \(K_2\), the noxious weed is effectively kept under control by the benign one. If the environment is slowly changing such that \(K_2\) slowly increases each season (from \(K_2(0)\) to \(K_2(2)\) to \(K_2(3)\) in Fig. 2), assuming a relatively low competition coefficient of 0.9, the results in Fig. 2a would be expected, that is, a slow increase in the noxious weed each year along with a decrease in the benign one. However, if the competition coefficients are much larger, say 1.1, the change in the weed community, when it occurs, will be completely abrupt (as in Fig. 2b). Furthermore, once the switch to a complete dominance of the noxious weed is complete, if we engineer a reversal of the environmental change, the return to a dominance of the benign weed may occur at a far lower value of the carrying capacity of the noxious weed than when we moved in the other direction. This effect is commonly referred to as hysteresis, and is part of an overall structure known as a cusp catastrophe (see Fig. 3). This arrangement of regime shift based on the value of the competition coefficients means that there is unlikely to be a “signature” in any available data that could serve as a warning that a regime shift was on the horizon. It could be argued (Scheffer,
2010) that one might find a “flickering” of the system if there were enough variability in the environment surrounding the system, such that the competitors existed either on one side of the cusp or the other, or a bimodal distribution could serve as an indication that a regime shift to the noxious weed was possible. Additionally, as noted by Van Nes and Scheffer (2007), there will be a generalized “critical slowing down” (as can be seen in Fig. 2b) as the catastrophic bifurcation is approached. However, such signals may not be observable under real world conditions. Reflecting on Fig. 2b, it is clearly observable that there is a critical slowing down, but what a farmer or farm manager is likely to confidently observe is the qualitative fact that there is a dramatic yet temporary increase in the noxious weed at the beginning of the season (indeed the size of that increase may even decrease from year to year, as it does in Fig. 2b, resulting in a false sense of security), but fail to be able to detect the more subtle signal of critical slowing down. The formality of this structure, as reflected in Figs. 2 and 3, suggests that the transformation could easily happen with no forewarning at all.

3. Basin boundary collisions with and without environmental drivers

The fundamental structure of the phenomenon of a basin boundary collision (Vandermeer and Yodzis, 1999) suggests the likelihood of regime change without prior signal, generally. The extensive literature on regime change has recognized this issue as a general phenomenon (e.g., Hastings and Wysham, 2010; Scheffer, 2009; Scheffer et al., 2009). Here I provide two examples that are particularly relevant to agroecosystems.

Consider a group of small independent producers: First presume that their decision as to how much land to put into the production of a particular crop is informed only by the present price of the crop; second, presume that excessive intensity of production may lead to some ecological damage that would translate into a loss of expected yield. Given these two assumptions, it can be argued (Vandermeer, 1997) that two alternative states emerge, one with high product prices, market volatility, ecological decline, and low production, the other with low product prices, stable markets, ecologically benign and high production. One analytic formulation of this basic idea is through a composed iterative map, using the variables, price, total production, and yield. Using production as the salient state variable, and composing the functions relating (1) production to cultivation intensity, (2) cultivation intensity to unit commodity price, and (3) unit commodity price to production, production at time \( t+1 \) can be represented as a quartic map,

\[
P_{t+1} = [abnP_t(1 - nP_t)][a - aNP_t(1 - nP_t)]
\]

where \( a, b, \) and \( n \) are constants associated with each of the contributing iterative map functions, “\( a \)” referring to the rate at which cultivation intensity will increase with increasing product price, “\( b \)” referring to the increase in regional yield with increasing cultivation intensity, and \( n \) referring to the rate of decline in prices with increasing regional yield (Vandermeer, 1997).

Using this model framework, many scenarios can be generated, three of which are illustrated in Fig. 4. First (Fig. 4a), with no environmental forcing at all (the model parameters remain constant over time), it is possible that the system will remain in the high production/low volatility mode for a long period of time, but then suddenly, and without warning jump to the low production/high volatility mode, where it stays. Note that this happens with no change in environmental or market conditions at all. It is a regime change, to be sure, but one in which there is no discernable signal that the regime will change, a clear surprise for anyone living for a long period of time within the first mode.

Second (Fig. 4b), a small secular change has been introduced, such that the underlying conditions change gradually through time. The change could represent a gradually increase in market expectations, or a gradual decline in ecological conditions, or a variety of other possibilities. The point is not the particular driver of the change, but the fact that a slow secular change is happening. At around time 300 the system dramatically jumps from one syndrome to the other. The only clue that a jump was imminent, was a gradual increase in the amplitude of the fluctuations, yet even that was not really a clue since the particular amplitude at which the transition would occur is not predictable. Worse is the regime change at approximately 650 time units, in which a very stable situation is suddenly interrupted by a transition to a highly volatile situation. In both cases, the transition that resulted from
the slow secular change in underlying parameters was sudden and surprising.

Third, (Fig. 4c) a small secular change similar to the case of Fig. 4b generates what might be detectable behavior before the full transition to the volatile state. The so-called “flickering” (Scheffer, 2010), in which the system alternatively visits two alternative sites unpredictably, appears to occur just before a full emersion into the highly volatile behavior. It is a matter of opinion whether there is significant flickering between about times 200 and 320, or whether that flickering continues through about time 450. Such details are not the concern of the present work. The possible existence of flickering suggests that a signal of impending regime change might be detectable.

Thus, in the examples of Fig. 4 we have two cases of the inevitability of surprise, one in which no external driving force at all is involved (Fig. 4a) and one in which a slow secular change is involved (Fig. 4b). The third example (Fig. 4c), using a framework very similar to the previous example (Fig. 4b), it is at least conceivable to detect a signal of the impending regime change, although detecting such a subtle signal may be, in practice, difficult.

The model giving this result is a two-humped iterative map, as illustrated in Fig. 5, the derivation of which is fully described elsewhere (Vandermeer, 1997). In Fig. 5a, the parameters are set such that there two chaotic attractors, representing the alternate “syndromes of production.” Each attractor is constrained within a basin of attraction and the system will remain in whichever of the two basins it began, in perpetuity. In Fig. 5b, there is a “ghost” attractor where the former upper attractor existed, and even though the system may start in the zone of the ghost attractor, it will inevitably jump into the alternative attractor at some point. This is a classic example of a basin boundary collision (Vandermeer and Yodzis, 1999) in which the edge of what was the upper chaotic attractor in Fig. 5a, collides with the boundary of the basin of the lower chaotic attractor. In Fig. 5c, the system is at what is effectively a double basin boundary collision in which both basins of attraction are intercepted by the edges of both chaotic attractors. In Fig. 5d the two previously separate attractors have fully penetrated each other’s basins and unpredictably wander over the whole range of extreme values. The scenarios of Fig. 4 can easily be seen with reference to the various panels of Fig. 5. The scenario in Fig. 4a is precisely an example of the first basin boundary collision pictured in Fig. 5b. The scenario in Fig. 4b is somewhat more complicated in that first (around time = 300), the changing parameter moves the system from a situation like that in Fig. 5a, to a system similar to Fig. 5b, except with the inverse basin boundary collision, and then with further change in parameter the other basin boundary collision occurs, putting the system into a state similar to that portrayed in Fig. 5d.

4. Global bifurcation and resource categories

Frequently in a case of biological control using a specific natural enemy, the control agent itself will have alternative forms of control exerted on its population. Thus, for example, insectivorous birds might be significant predators in an agroecosystem, but the effective control over their populations may be through the limitation of nesting sites, rather than food sources. Or, parasitoid wasps might be limited by the availability of flowering resources rather than the insect pests they parasitize. In these and many other cases that could be cited, the controlling resources are of two distinct qualities (Vandermeer and King, 2010). On the one hand (e.g., nesting resources) the resource itself is not affected by the predator. These resources are called “fixed”. On the other hand (e.g., the prey or host), the resource is indeed affected by the predator. These resources are called “variable.” The dynamics of a natural enemy under these two constraints can be modeled with
the following pair of equations (Vandermeer and King, 2010)

\[
\frac{dC}{dt} = a \left( \frac{K - C}{K} \right) CR - mC \tag{3a}
\]

\[
\frac{dR}{dt} = rR(1 - R) - c(\frac{K - C}{K}) CR \tag{3b}
\]

where \(C\) is the consumer (the potential natural enemy) and \(R\) is the resource (the potential pest), \(a\) is the consumption rate, \(c\) the conversion rate, \(r\) the intrinsic rate of the resource and \(K\) the carrying capacity of the consumer. The fixed resource is modeled as the carrying capacity of the consumer while the variable resource is the second dynamic equation (see complete derivation in Vandermeer and King, 2010).

These equations lead to the result that under certain conditions the predator (consumer) will be mainly held under control by the fixed resource and under other conditions by the variable resource. There is a particular point at which the dynamic rules of the system are close to a global bifurcation in which the system moves from one to the other underlying dynamic framework. The consequences of this underlying dynamic are illustrated in Fig. 6. In this example, the pest population has a population flush each spring, followed by a flush of the natural enemy (not shown in the figure) which drives the pest to near zero, well below the economic threshold (horizontal dashed line in Fig. 6). This happens four years in a row (and perhaps many years previous to these particular four), but then a very small change in a parameter in the model, any parameter, shifts the system in season 5, and, after that first population flush, which seems identical to the previous years, the pest flushes well above the economic threshold. In Fig. 6 the particular parameter that was changed was \(r\), but a similar pattern could have been generated by changing any of the parameters.

This example is particularly interesting in that it represents an unusual form of bifurcation (Vandermeer and King, 2010), a global bifurcation in which the vector field itself is dramatically transformed such that the basins of attraction for the two alternative equilibrium points are themselves transformed, as illustrated in Fig. 7. The basin boundary collision of the previous section effectivley represents a situation in which a chaotic attractor becomes large enough to transcend the boundary of its original basin, yet the overall structure of the vector field within the basins remains qualitatively the same. The bifurcation discussed in this section, the global bifurcation, is a qualitatively distinct transformation of the vector field itself. The switch from Fig. 7b to either Fig. 7a or c represents a change in the overall organization of the vector field, such that beginning with a small number of natural enemies, the system can go either to control by the natural enemy (Fig. 7a) or control from the alternative resource (Fig. 7b). When the system is close to the bifurcation point (Fig. 7b), a change in any parameter of the system can distort the dynamics in either direction. Thus, a system that repeatedly moves toward control from the variable resource (Fig. 7a or the first four seasons of Fig. 6) can easily be pushed into control from the fixed resource (Fig. 7c or the last two seasons of Fig. 6), by the slight modification of any of the parameters in the system (Eqs. (3a) and (3b)). And the emergent behavior of the system will give any a clue as to the impending dramatic change in the regime.

5. Discussion

It is an undeniable fact that ecosystems sometimes, perhaps frequently, undergo dramatic changes, what has come to be called “regime change” in recent literature (Kéfi et al., 2007; Rietkerk et al., 2004; Scheffer, 2009), usually associated with some sort of alternative equilibrium structure (May, 1977). Along with the acknowledgement that such behavior is not only possible, but likely, there has recently been growing interest in developing techniques which can be used to detect some sort of signal that a system was about to undergo a regime change (Scheffer et al., 2009; Scheffer, 2009). Such a goal has obvious importance in any ecosystem that provides services to humans, an especially important one of which is the agroecosystem.

However, it is perhaps a short step from the search for a clue to the assumption that such a clue necessarily exists. Here I elaborate on several well-known theoretical structures that suggest the complete absence of any sort of practically detectable signal. These theoretical structures are generally well-known and appreciated in the growing literature on regime change (e.g., Carpenter and Brock, 2010; Dakos et al., 2011). In the light of that recognition, I argue in this paper that advanced monitoring for important regime changes should not be the exclusive goal of managing with environmental change in mind. Although recognizing such signals should be a top priority, older philosophical principles that evolved under the assumption of an underlying uncertainty, such as the “precautionary principle,” should not be abandoned. Indeed, in this paper I provide examples of a variety of theoretical structures that suggest caution should be the byword in ecosystem management, even if no evident signal of impending regime change is evident.
Historical precedence for this point of view can be found in the economics literature, where it is arguably the case (Skidelsky, 2009) that the fundamental Keynesian paradigm is founded on a recognition of the difference between risk and uncertainty. Risk can be managed, similarly to ecosystem changes that we can foresee through careful statistical procedures. But uncertainty signifies the need to proceed with caution, in both economics and ecology.

The admonition to proceed with caution has been common in practical applications of technology to ecosystems, canonized in the phrase “the precautionary principle,” one popular definition of which is,

In order to protect the environment, the precautionary approach shall be widely applied by States according to their capabilities. Where there are threats of serious or irreversible damage, lack of full scientific certainty shall not be used as a reason for postponing cost-effective measures to prevent environmental degradation.

This definition is principle 15 of the Rio Declaration of the 1992 United Nations “Conference on Environment and Development” (UNCED), frequently referred to as the Earth Summit. Although the principle is usually thought of as a political/policy issue, the argument in this paper is that there are solid reasons stemming from ecological theory to advocate the precautionary principle. If an alternative basin of attraction can be plausibly identified and if that alternative basin suggests a “threat of serious or irreversible damage,” the underlying dynamics of the system might very well not generate any caution sign before that undesirable regime change becomes manifest, and “cost-effective measures” should be employed to avert slipping into that alternative basin. Waiting for a signal that the threat is imminent may very well be a fool’s errand, given what we now know about ecological dynamics.

References