III ANALYSIS OF REGIME SHIFTS: ROLES OF STOCHASTICITY AND EXPERIMENTS

Introduction

Statistical methods are powerful for analyzing frequent events, but weak for analyzing uncommon events such as regime shifts. The rarity of regime shifts poses deep problems for both understanding and prediction. For any given regime shift, many different models are equally appropriate. There are straightforward statistical methods for sorting among rival models, but these depend on observing many regime shifts. Without extensive data, it may be difficult to decide which model is most appropriate. Consequently, there is considerable doubt about our ability to predict future regime shifts on the basis of any particular model.

This chapter addresses some statistical problems posed by regime shifts, using both models and data. A minimal model, which represents no particular ecosystem but is transparent and easy to understand, is used to explore the detectability of alternate equilibria using time series data. Then, an example using field data is introduced. This example illustrates the use of time-series data and comparative data together to estimate parameters for regime shifts.

There are three parts to the chapter. First, a graphical model of regime shifts is used to explain some of the implications of experimental manipulations and sources of error that affect the detection of regime shifts. Second, a minimal model is presented for drawing inferences about regime shifts from time series of ecosystems. The model is designed to address issues such as duration of time series, ability to perform informative experiments, the effects of stochastic shocks, and the effects of observation error. This minimal model suggests that it is difficult to characterize regime shifts. In particular, there is considerable risk that analysis of a typical data set will conclude incorrectly that the system has only a single stable point. The third part of the chapter presents an analysis of a well-studied ecosystem, Lake Mendota, to demonstrate the challenges of assessing regime shifts for a real data set. The case for regime shifts in the eutrophication of Lake Mendota depends not only on time series, but also on detailed process measurements, a comparative multi-lake data set, and historical information.

Detecting Regime Shifts: A Graphical Model

Ball-and-cup diagrams (Fig. 14) have often been used to describe alternate stable equilibria of ecosystems. Alternate stable states are one of many potential causes of regime shifts, as described in Chapter I. In the ball-and-cup diagram, the topography represents slowly changing variables (Fig. 14A). The basins are stable points, and the hilltops are thresholds. For simplicity, only two basins and one threshold will be shown here, but any number of basins and thresholds are possible. The ball represents the level of an ecosystem variable that changes more rapidly than the topography changes. In the case of lake eutrophication, the amount of phosphorus in the lake sediments is represented by the topography, and the water quality (turbidity, phosphorus in the water, or biomass of phytoplankton) is represented by the ball (Chapter II). Shifts among equilibria can be caused by external shocks that move the ball from one basin to another (Fig. 14B). The shock can be a stochastic event, such as a violent storm, or a deliberate manipulation caused by an experiment or a management action. Changes in slow variables can alter the topography and shift the ecosystem to another basin (Fig. 14C). For example, slow buildup of phosphorus in sediments of a lake shrinks the basin of the clear-water regime and moves the lake toward the turbid-water regime (Chapter II).

Now consider an observer who wishes to determine whether the ball-and-cup system has one basin or two. The observer measures the rapidly-changing variable, which corresponds to the horizontal position of the ball (Fig. 15). This measurement and its uncertainty are represented by the horizontal error bars in Fig. 15. Observation error (the standard deviation of measurements of the position of the center of the ellipsoid) increases the length of the error bar. If the center of the ellipsoid is in only one basin throughout the period of observation, and the observation error is smaller than the basin, the observer may conclude that the system lies in a rather narrow range (Fig. 15A). On the other hand, if the center of the ellipsoid stays in one basin but the observation error is larger than the basin, the observer may conclude that the system occupies a broad range (Fig. 15B). Although the observations may extend to two or more basins, the observer would be uncertain about the number of basins.

Experiments make it easier to perceive the number of basins, especially if the observation error is smaller than the basin (Fig. 15C). The experimental perturbations lead to observations that are clustered, and the clusters correspond roughly to the basins. If natural events that can cause regime shifts are rare, experiments may be a rapid and effective way to determine the number of regimes.

Process variance is the term used for the variance shocks that affect the location of the ball. If the center of the ball lies in one basin but the ball is subject to small random disturbances, the observer will perceive some variability in the location of the ball. However, this variability is not created by measurement errors but instead it is caused by the small random changes in the location of the ellipsoid. These variations are called process variance, or process errors, to distinguish them from observation errors (Hilborn and Mangel 1997). The distinction between process error and observation error will be made in a mathematical model in the next section.

Process variance can make it easier to perceive the number of basins (Fig. 15D). Essentially, the random shocks act as “natural experiments”. The observer perceives clusters of observations, which correspond roughly to the locations of the basins. For the observer to perceive the two basins, the frequency of shifts between basins must be not too fast and not too slow. If the shifts between basins are very frequent, then it does not matter that there are two basins. The system acts as if the topography was flat. If the shifts between basins are very infrequent, then we have the same situation as in Fig. 15A and the observer perceives only one basin.

In summary, this pictorial model captures the essential ideas of regime shifts in a simple way. The topography represents the slowly-changing ecosystem variables and the position of the ball represents the rapidly-changing ones. The basins are the different regimes and the hilltops are the thresholds. We imagine a human observer who wishes to know how many regimes exist in the system using measurements of the position of the ball. Observation errors, or noisy measurements, blur the location of the ball. Random disturbances, or process errors, change the location of the ball and therefore cause variability in measurements. Experiments can significantly improve the observer's ability to perceive the number of basins. Process errors can also improve perceptions, if disturbances that shift the system between basins are not too frequent and not too rare. In the next section, we will explore these ideas further using a mathematical model.

Prospects for Identifying Alternate Equilibria: A Simulation

In ecological modeling, the term "minimal model" is used for models that are just complex enough to represent the abstraction under study, but no more complex than necessary (Scheffer and Beets 1994). In this section, a minimal model of a system with alternate equilibria is used to study the effects of observation error and experimentation on the ability to detect alternate equilibria. Like the pictorial model of the previous section, this minimal model includes a simple one-dimensional representation of an ecosystem with alternate equilibria subject to process error, and an observer who samples the system with varying degrees of observation error. The observer can also perturb the system experimentally. The minimal model is mathematical, so it can represent the processes of gathering and analyzing data for a variable system subject to alternate equilibria. Using this model, I examined the probability of correctly detecting alternate equilibria with and without experimental manipulation, with different magnitudes of observation error.

A Minimal Model for Studying Alternate Equilibria in Time Series

This section presents a mathematical model that will be used to explore possibilities for detecting alternate equilibria in time series. The model is as simple as possible, yet contains key features of ecosystem data relevant to the analysis of regime shifts. Thus the model includes a dynamic system with alternate equilibria and process variance, a sampling process with observation error, and an analytical procedure for inferring the number of equilibria. This model is an arbitrary, deliberately simplified representation of a time-varying system with two stable equilibria. The simplicity of the model is useful for demonstrating some general points about detecting alternate equilibria, even though the model does not represent any particular ecosystem. The model first appeared in Carpenter (2001).

The model assumes that time series, of varying length, can be observed for the ecosystem of interest. The ecosystem is subject to stochastic shocks, similar to the effects of climatic variability on ecosystems. The observations are subject to measurement error, as occurs in actual ecological data. In some simulations it is possible to perform experiments intended to switch the ecosystem between regimes. The simulated ecosystem has two different equilibria and an unstable threshold, but this fact is unknown to the investigator. Instead, the investigator must evaluate statistical evidence for the regime shifts. Artificial data sets are generated using the ecosystem model. Then, these artificial data are analyzed for evidence of alternate equilibria (Fig. 16). This model uses a large number of simulations to measure the probability of detecting alternate equilibria when in fact alternate equilibria are the true situation.

The simple ecosystem model is as follows. Time series data sets of an ecosystem variate X were calculated by repeated simulations with the difference equation

$$X_{t+1} = X_t - \beta_0 (X_t - \beta_1)(X_t^2 - \beta_2) + v_t \qquad (1)$$

The β's are parameters and v represents random shocks to the system, or process errors (Hilborn and Mangel 1997). These process errors were sampled from a normal distribution with mean zero and standard deviation s_v.

Equilibria are calculated by setting $X_{t+1} = X_t = X^*$ and solving for the equilibria X^*. If $\beta_2<0$, there is one stable real equilibrium ($X^* = \beta_1$) and two complex equilibria. The complex equilibria are of no ecological significance, and can be ignored. If $\beta_2>0$, there are three real equilibria (β_1, $\beta_2^{0.5}$, and $-(\beta_2)^{0.5}$). The largest and smallest of these three are attractors, and the intermediate one is a repeller.

Measurements of ecological time series data are of course imperfect, so it is desirable to include observation errors in the simulation exercise. Observations of X, symbolized by Y, were simulated as

$$Y_t = X_t + w_t \tag{2}$$

where w_t is observation error (normally distributed with mean zero and standard deviation s_w).

For simulations shown in this chapter, parameter values were $\beta_0 = 0.1$, $\beta_1 = 1$, $\beta_2 = 0.16$, $s_w = 0.1$ or 0.0001, and $s_v = 0.05$. Thus there were two stable points at $X = -0.4$ and $X = 1$, and one unstable point at $X = 0.4$.

Two types of data sets were generated, non-experimental data and experimental data. To generate non-experimental data, simulations of a given number of time steps were calculated from a randomly-chosen starting point on the interval between the two stable points. Duration of the time series was varied to represent different durations of long-term ecological studies. For experimental data, an infinite number of experimental designs can be imagined, and one approach was chosen arbitrarily for this book. Trials with several different experimental designs indicated qualitatively similar results in all cases. To generate experimental data, simulations were re-started at a new random starting point every 12 time steps. This represents manipulations that regularly re-set the initial conditions of the ecosystem. For both non-experimental and experimental scenarios, 1000 independent data sets were simulated.

The probability of identifying alternate equilibria was calculated for each simulated time series. Time series were fit to a cubic polynomial

$$Y_{t+1} = b_0 + (1 + b_1)\, Y_t + b_2\, Y_t^2 + b_3\, Y_t^3 + \varepsilon_t \qquad (3)$$

The b_i are the regression parameters. ε is the model error, which is assumed to be normally distributed with mean zero and variance to be estimated from the data. Note that Equation 1 expands into a cubic polynomial like Equation 3. The estimated parameters b_i are functions of the unknown parameters β_i, and the observed model error ε is a function of the unknown process and observation variances.

To account for observation error, the entire trajectory was fitted by least squares to Equation 3 (Hilborn and Mangel 1997). That is, for a given set of parameter values, the entire time series was simulated starting with the first observed value. Then the sum of squared differences between simulated and observed time series was calculated. The optimal parameter values were taken to be those that minimized this sum of squared differences.

Once the parameter estimates were calculated, the equilibria ($Y_{t+1} = Y_t$) were calculated as the roots of the polynomial

$$0 = b_0 + b_1 Y_t + b_2 Y_t^2 + b_3 Y_t^3 \qquad (4)$$

This polynomial has either one real root and two complex roots (indicating one stable equilibrium), or three real roots (indicating three equilibria). In the latter case only two of the equilibria are stable. The probability of detecting alternate equilibria was calculated as the proportion of simulated data sets that yielded three real roots.

Detection of Alternate Equilibria

In the time series simulated without experimental manipulation, the probability of detecting alternate equilibria was relatively low (Fig. 17A). In the absence of manipulation, the system is not likely to leave the initial regime over the course of a simulation. Therefore, the analysis is likely to detect only one of the two stable equilibria. In simulations with larger observation error, the probability of detecting both stable equilibria is near 0.5 and does not change perceptibly as sample size increases. This phenomenon was illustrated by the ball-and-cup diagrams with large observation error. Errors in measuring the value of X can be interpreted as shifts between equilibria.

In simulations with low observation error, the probability of detecting both stable equilibria is about 0.4 for small sample sizes, and declines to about 0.25 for large sample sizes. This decline occurs because parameter estimates are increasingly precise for large sample sizes, but the system occupies only one stable domain. There is an increasingly precise characterization of a narrow range of the system's potential dynamics. This leads to the conclusion that the system has only one stable equilibrium, because the other stable equilibrium is unlikely to occur in any given time series.

In the time series simulated with experimental manipulation, alternate equilibria are more likely to be detected (Fig. 17B). Each time the system is reset experimentally, there is a chance that it will be moved to a different domain of attraction. With low observation error, the probability of detecting alternate equilibria climbs above 0.9 for the longer time series. As the time series get longer, one observes more switches between domains of attraction and parameter estimates become more precise. With high observation error, however, the probability of detecting alternate equilibria hovers between 0.5 and 0.6, and actually appears to decline slightly for longer simulations. Even though experimental manipulation switches the system between domains of attraction, and more switches occur in longer time series, the high observation error makes it difficult to determine that switches have occurred.

As was illustrated using the ball-and-cup model, certain magnitudes of process error have an effect comparable to experimental manipulation (Carpenter 2001). The random shocks of process error can switch the system between stable domains, thereby increasing the probability that alternate equilibria will be detected. If process errors are too small, the shocks will not switch the system between stability domains and therefore will not increase the probability of detecting alternate equilibria. If process errors are too large, the system can jump between stability domains frequently. If the shift between stable equilibria occurs frequently due to process error, then the probability of detecting alternate equilibria will be low. Of course, such alternate equilibria may have no ecological significance. A system that rapidly switches between alternate equilibria due to process error is simply noisy. It does not settle into any regime for more than a few time steps. Even though there are multiple equilibria, they have no practical significance for the dynamics of the system.

Despite its simplicity, this exercise illustrates important points about detecting alternate equilibria. The ecosystem must switch between equilibria during the period of study. Ideally several regime shifts should be observed. Regime shifts can be produced by natural variability (represented by process error in this model) or by experimental manipulation. The data must be high quality. Observation error obscures the actual state of the system, and increases the difficulty of detecting alternate equilibria. Overall, this exercise suggests caution in using time series data alone to conclude whether or not alternate equilibria exist. The statistical power to detect alternate equilibria will be low for most ecological time series.

In reality, the ecologist searching for evidence of alternate equilibria faces many difficulties that are not represented by the simple exercise presented here. In this exercise, the investigator knows the true system structure. The choice between fitted models is a simple dichotomy between a single stable equilibrium and double stable equilibria. Thus, detection of alternate equilibria hinges entirely on estimating the parameters correctly. These simplifications are useful for exposing some of the problems in detecting alternate equilibria, which are actually made more severe by the complications of studying real ecosystems. In actual research, there are many plausible models, the set of models may or may not include a good approximation to the true system structure, and many parameters are unknown and must be estimated from data. In ecosystem studies, conclusions about the *absence* of alternate states must always be suspect. It is plausible that the observations are simply not up to the task.

Multiple Lines of Evidence: A Case Study

This third and final section of the chapter considers the detection of alternate states in field measurements from ecosystems. The previous section showed that it is difficult to detect multiple equilibria from time series data alone. The most convincing examples of multiple states in ecosystems rest on multiple lines of evidence, including long-term observations, experiments, and comparative studies (Carpenter 2001). This section will combine long-term and comparative data to explore features of alternate equilibria in an eutrophic lake.

To illustrate how several kinds of evidence can be used to assess the possibility of multiple equilibria, I will consider the eutrophication of Lake Mendota (Wisconsin, USA). For this lake, relatively long time series data are available, and a number of measurements have been made of key process rates. Also, some large management manipulations have been performed in attempts to improve water quality (Carpenter et al. 2004). These manipulations may increase the probability of detecting transitions between states in Lake Mendota, if such transitions have occurred. It turns out that comparative studies (measurements made on a large number of similar lakes) are important for the analysis of Lake Mendota.

Lake Mendota is among the world's best-studied lakes (Brock 1985, Kitchell 1992, Carpenter et al. 2004). It is culturally eutrophic due to excessive phosphorus inputs from its watershed (Soranno et al. 1996, Bennett et al. 1999). The eutrophic state is stabilized by phosphorus accumulated in watershed soils, sediments of tributary streams, and sediments of the lake itself (Bennett et al. 1999, Reed-Anderson et al. 2000). The hypolimnion of Lake Mendota is anoxic for about 60 to 90 days each summer so there is significant potential for phosphorus recycling (Nürnberg 1995). Direct measurements of internal phosphorus recycling show that it can exceed inputs by several-fold (Stauffer 1974, Soranno et al. 1997). On the basis of these publications, the model of eutrophication presented in Chapter II seems plausible for Lake Mendota.

Do phosphorus dynamics in Lake Mendota correspond to qualitative patterns that are expected from the eutrophication model of Chapter II? To examine some correlations relevant to this question, I analyzed the phosphorus budget time series presented by Lathrop et al. (1998). These authors measured phosphorus budgets for Lake Mendota for 21 consecutive years. The lake's P budget can be described by the equation

$$P_{t+1} = P_t + L_t + R_t - S_t - X_t \qquad (5)$$

The subscript denotes year, which starts on 15 April of each calendar year in this data set (Lathrop et al. 1998). P is mass of phosphorus in the lake, L is annual input of phosphorus from the watershed, R is recycling from sediments, S is sedimentation and X is export in surface water flowing out of the lake. All quantities are measured in kg. Lathrop et al. (1998) directly measured P, L and X, with P measured on 15 April each year and L and X measured for the annual cycle beginning 15 April. Thus net recycling ($R_t - S_t$) can be calculated by difference. In addition, average summer P mass was estimated using data from Lathrop and Carpenter (1992) and unpublished data of the North Temperate Lakes LTER site (http://lter.limnology.wisc.edu).

If the eutrophication model of Chapter II is appropriate, we would expect P content of the lake water to be directly related to recycling. Four indicators of recycling were examined for correlations with summer P mass in the lake (Fig. 18). Annual change in P mass ($P_{t+1} - P_t$) is directly related to summer P mass, as expected (Fig. 18 A). Net recycling ($R_t - S_t = P_{t+1} - P_t - L_t + X_t$) is also directly related to summer P mass (Fig. 18 B). Two additional indices of recycling can be calculated from residuals. Residuals from the regression of the quantity ($P_{t+1} + X_t$) versus L_t represent the amount of P accumulated through the year that cannot be explained by inputs. P recycling is the likely mechanism for such accumulation. These residuals are directly correlated with summer P mass (Fig. 18 C). An alternative indicator of recycling is the regression of the quantity (P_{t+1}- P_t + X_t) versus L_t. These residuals should measure the increment of P mass during the year that cannot be explained by load. This surrogate for recycling is also directly correlated with P_t.

All of the rough indicators of recycling shown in Fig. 18 are directly correlated with P mass in the overlying water during summer. These results are consistent with the feedbacks invoked in the regime-shift model for lake eutrophication (Chapter II). According to this model, high P mass in the water during summer increases the duration and spatial extent of anoxia in the hypolimnion of the lake. Anoxic conditions increase the solubility of phosphate, thereby increasing the flux of phosphate from sediments to the overlying water.

While these budget analyses are consistent with the eutrophication model of Chapter II, they suffer from the problem that phosphorus in the water column may turn over many times during the year. For example, Soranno et al. (1997) showed that recycling and sedimentation exceed annual inputs, suggesting that each phosphorus atom cycles several times a year between sediments and the overlying water. Thus the measures of net recycling from budget data may miss a substantial amount of phosphorus flux with significant implications for eutrophication.

Ludwig et al. (2003) addressed the fast dynamics of sedimentation and recycling by using a continuous time model, which was then integrated to calculate the annual fluxes. Their continuous time model for dynamics of P in the water column is

$$dP/dt = L - (s + h) P + r [P^q / (m^q + P^q)] \qquad (6)$$

Here L is the continuous input rate, s and h are rate constants for sedimentation and hydrologic outflow, respectively, r is the maximum recycling rate, and m is the P mass in the water for which recycling is half the maximum rate. The parameter q sets the steepness of the recycling versus P curve when $P \approx m$. To integrate this differential equation over one year, Ludwig et al. (2003) assume that the recycling term (in square brackets) is approximately constant over the year. The resulting difference equation is

$$P_{t+1} = c P_t + [(1 - c) / (s + h)] \{L + r [P_t^q / (m^q + P_t^q)]\} \qquad (7)$$

where $c = \exp(-s - h)$. While this difference equation is a rough approximation, it does accommodate the fact that annual P flux to sediments may be substantially greater than the average mass of P in the lake, because the mass of P in the lake is the result of a rapid dynamic equilibrium between sedimentation and recycling.

For the purposes of this chapter, I focused on the recycling parameter r. If there is a threshold for eutrophication, it will depend on r. Measurements of r have been made in many lakes (Nürnberg 1984) so r is a good choice for illustrating the benefits of combining information from different studies.

The unknown parameter r was estimated by fitting Equation 7 to data using Bayesian inverse modeling. Details of Bayesian inverse modeling are presented in the Appendix. It is one of the methods available for estimating parameters using data from two or more independent sources. By convention, the first source of data is summarized as a probability distribution called a prior distribution (Appendix).

The prior distribution for r was calculated from data published by Nürnberg (1984, her Table 6). She presents 14 lakes with anoxic hypolimnions for which the mean r is 5.11 g m^{-2} y^{-1} with a standard error of 0.68. It was also necessary to provide a prior distribution for the parameters s and m. Because these are unknown, I used a prior that was rather uninformative (i.e. had a high variance). For s, I used a mean of 2 with standard deviation 0.5. For m, I used a mean of equal to the maximum observed P value (2.3) with standard deviation half as large as the mean. These choices are arbitrary. However, the outcome was not sensitive to small changes in these values. The value of q was fixed at 8, near the value estimated by Carpenter et al. (1999b). The choice of q had only a weak effect on the outcome, over the range of q values from 4 to 20.

The second source of data was the time series data of Lathrop et al. (1998), which was described previously in this chapter. Using Bayesian inverse modeling (Appendix), the time series information was combined with the prior distribution to estimate probability distributions for r, s and m. Probability distributions that contain information from both sources of data are called posterior distributions (Appendix). The posterior distribution for the parameter r is most interesting, because it reflects the impact of the prior data from Nürnberg (1984).

One-year-ahead predictions from Equation 7 are similar to observations, although there is considerable scatter (Fig. 19; correlation of observations with one-year-ahead predictions = 0.61). Parameter estimates were s = 0.48, r = 5.1, m = 3.4, and $\sigma = 0.32$. While this equation fits the data, it appears to underestimate both sedimentation and recycling. The quantities c and (1 – c) / (s + h) are 0.51 and 0.72, respectively. These suggest that about half the standing stock of P and three quarters of the load plus recycling carry over from year to year. In contrast, sediment trap and entrainment studies suggest that the standing stock of P may turn over several times each summer due to sedimentation and recycling (Soranno et al. 1997). Annually-resolved data may simply be unable to represent the fast dynamics of these processes during the summer season.

The prior distribution for r (Nürnberg 1984) had a heavy impact on the fit of Equation 7 (Fig. 20). The distribution of r estimated by Nürnberg (1984) for 14 lakes is sharply defined. In contrast, the likelihood function based on time series for Lake Mendota alone is flat and diffuse. Consequently, the posterior distribution that combines the Lake Mendota data with Nürnberg's data is similar to Nürnberg's distribution. Nürnberg's estimate of r is 5.11 g m^{-2} y^{-1} and the posterior estimate was 5.07 g m^{-2} y^{-1}.

The threshold for eutrophication of Lake Mendota depends heavily on the estimated parameter r. These calculations show that the time series data for Lake Mendota are not very informative about r. A wide range of r values is possible (none are very probable), based on the time series alone. When time series data are combined with the comparative data from Nürnberg (1984), however, the probability distribution of r is considerably narrower. This more precise estimate of r will lead to a more precise estimate of the threshold. Because the calculation of the threshold is complicated, and is not needed to demonstrate the value of multiple sources of data, it will not be presented in this book. Interested readers will find a more complete analysis of the Lake Mendota time series, including probabilities of various types of dynamics, in Stow et al. (1997), Lathrop et al. (1998) and Carpenter et al. (1999b).

Summary

A simple pictorial ball-and-cup model illustrates the basic features of one explanation for regime shifts, alternate stable equilibria. The pictorial model also illustrates the effects of observation error, process error, and experiments on the perception of alternate equilibria. Observation error represents uncertainty in the measurements of the variables that indicate the state of an ecosystem. Process error represents the effects of external shocks that disturb the state of an ecosystem. Experiments are useful for testing the possibility of alternate equilibria because they can produce changes in ecosystem state that are large relative to observation and process errors. Also, it is often easier to infer the causes of ecosystem change when they are associated with a direct manipulation of a particular variable.

A minimal mathematical model was created to study the possibility of detecting alternate equilibria. The mathematical model represents the same situation illustrated by the ball-and-cup model. Results of the mathematical model show that it is difficult to identify and predict regime shifts using time series data alone. Experiments were helpful for revealing alternate equilibria, especially when observation errors were relatively small.

Regime shifts can often be understood using multiple lines of evidence. In lake eutrophication, for example, our understanding of transitions between the clear-water and turbid-water states derives from long-term observation, comparative study of large numbers of lakes, responses of water clarity to whole-lake manipulation of driving factors, and knowledge of the underlying biogeochemical mechanisms (Chapter II). A dynamic model consistent with these mechanisms appears to fit time series data from Lake Mendota. Fit of the model was improved substantially by including a comparative, multi-lake data set for a key recycling parameter. This analysis demonstrates one statistical method for combining data from multiple sources to analyze regime shifts. In general, models grounded in several types of evidence may be the most powerful method available for analyzing and understanding regime shifts.

Statistical methods are powerful for frequent events, but regime shifts are infrequent by definition. Regime shifts are therefore difficult to characterize statistically unless many events can be observed, and this requires long time series, whole-ecosystem experiments, observations of many similar systems subject to similar regime shifts, and careful process studies. In ecosystem studies, conclusions about the *absence* of alternate states must always be suspect. It is likely that the observations are simply not up to the task.

Figures

Figure 14. (A) Ball-and-cup model for an ecosystem with two stable equilibria. (B) An external disturbance shifts the system between states. (C) A change in the topography shifts the system between states. Arrow shows the uplift that causes the left basin to shrink, moving the ball to the right basin. (Original)

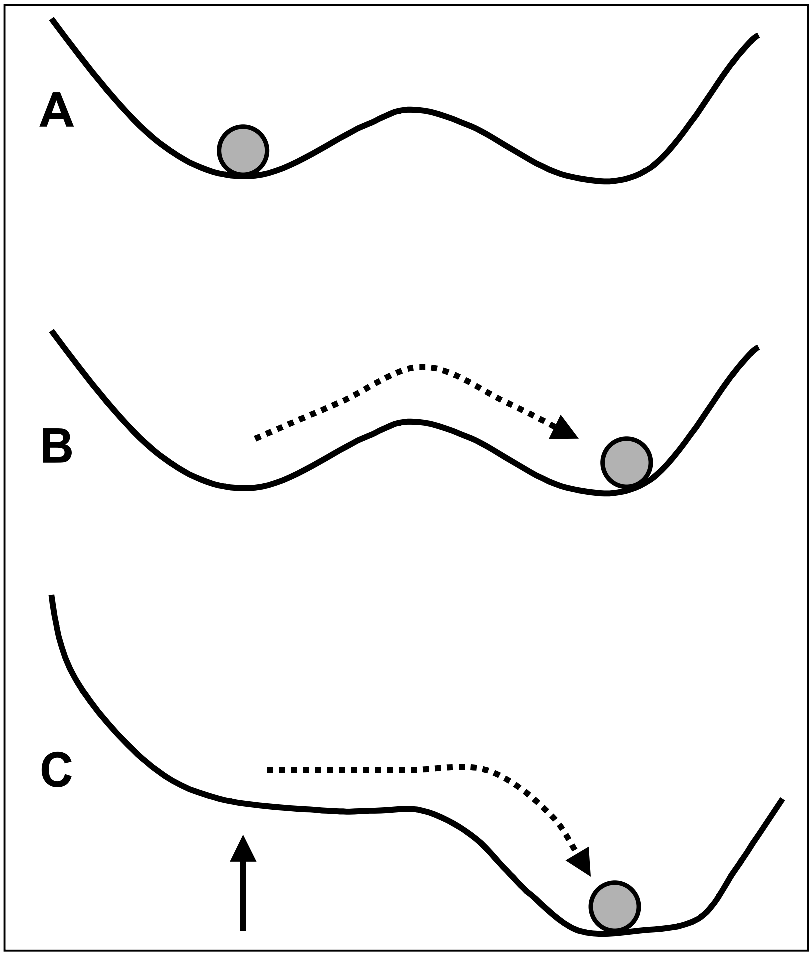

Figure 15. Ball-and-cup diagrams illustrating how observation error, experiments, and process errors can affect perception of the number of basins. The observer measures the horizontal position of the center of the ball. The perceived position of the ball is shown by the horizontal error bars. The black circles show the actual position of the ball. (A) Low observation error, with the ball in one basin throughout the period of measurement. (B) High observation error, with the ball in one basin throughout the period of measurement. (C) Low observation error, with experimental manipulation of the position of the ball. (D) Low observation error, with random shocks (high process variance) that disturb the position of the ball. (Original)

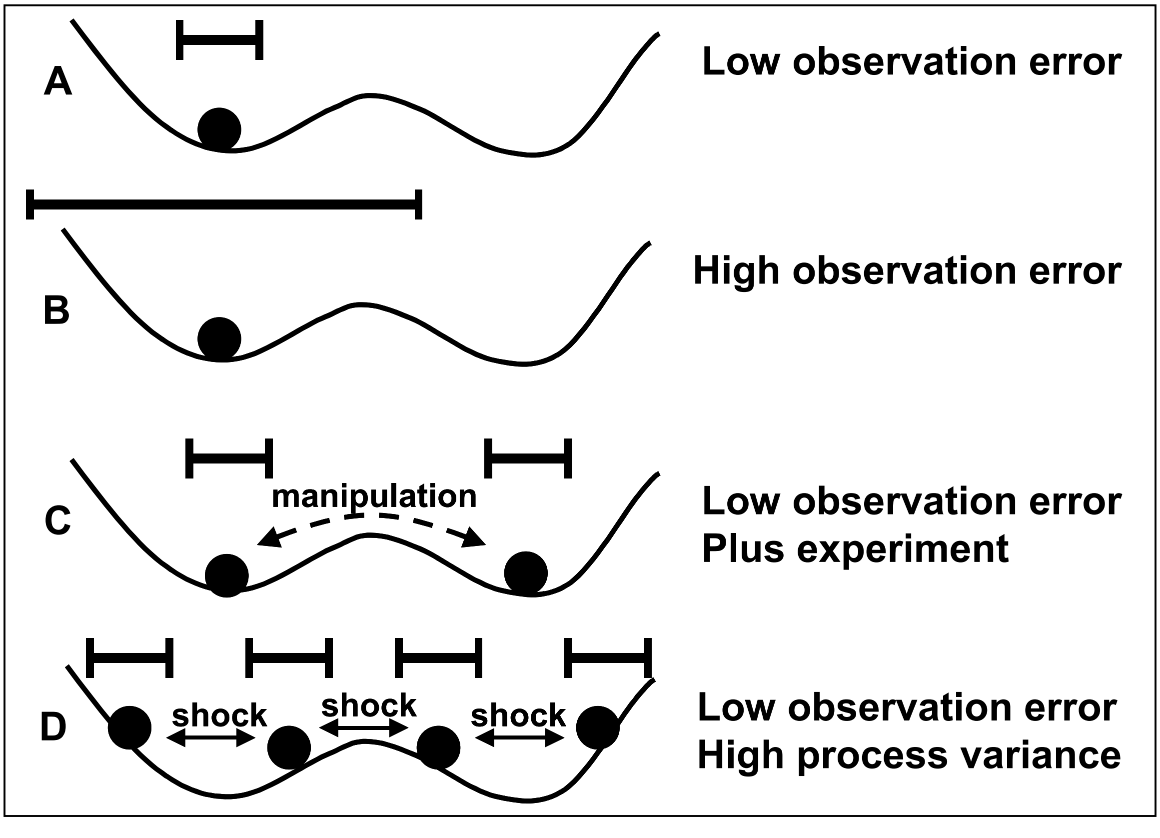

Figure 16. Flow chart of the simulation exercise for estimating the probability of detecting alternate equilibria. A. Simulations with no experimental manipulation. B. Simulations with experimental manipulation, in this case random resetting of the time series after 12 time steps. (Original)

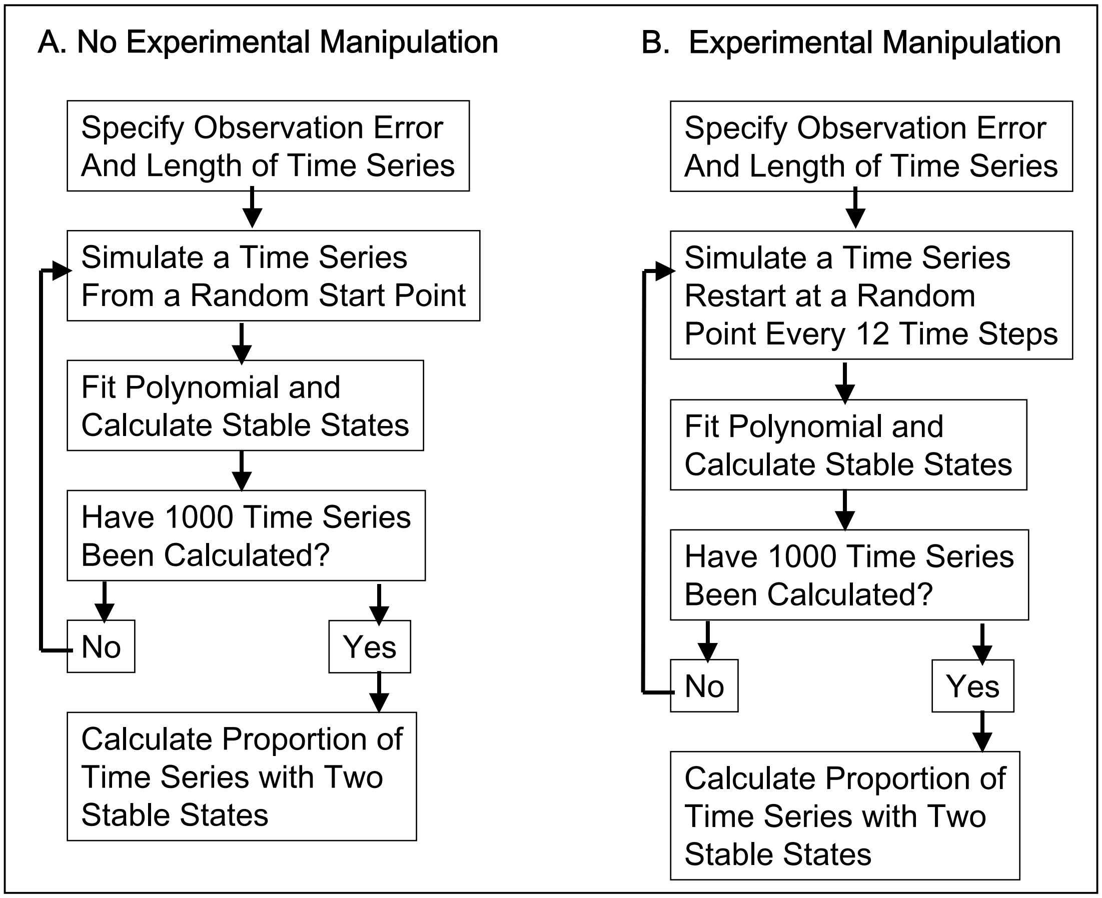

Figure 17. Proportion of 1000 simulations with two alternate stable equilibria versus number of time steps simulated, with observation standard deviation 0.1 (solid line) or 0.0001 (dashed line). A. Simulations without experimental manipulation. B. Simulations with experimental manipulation. (Original)

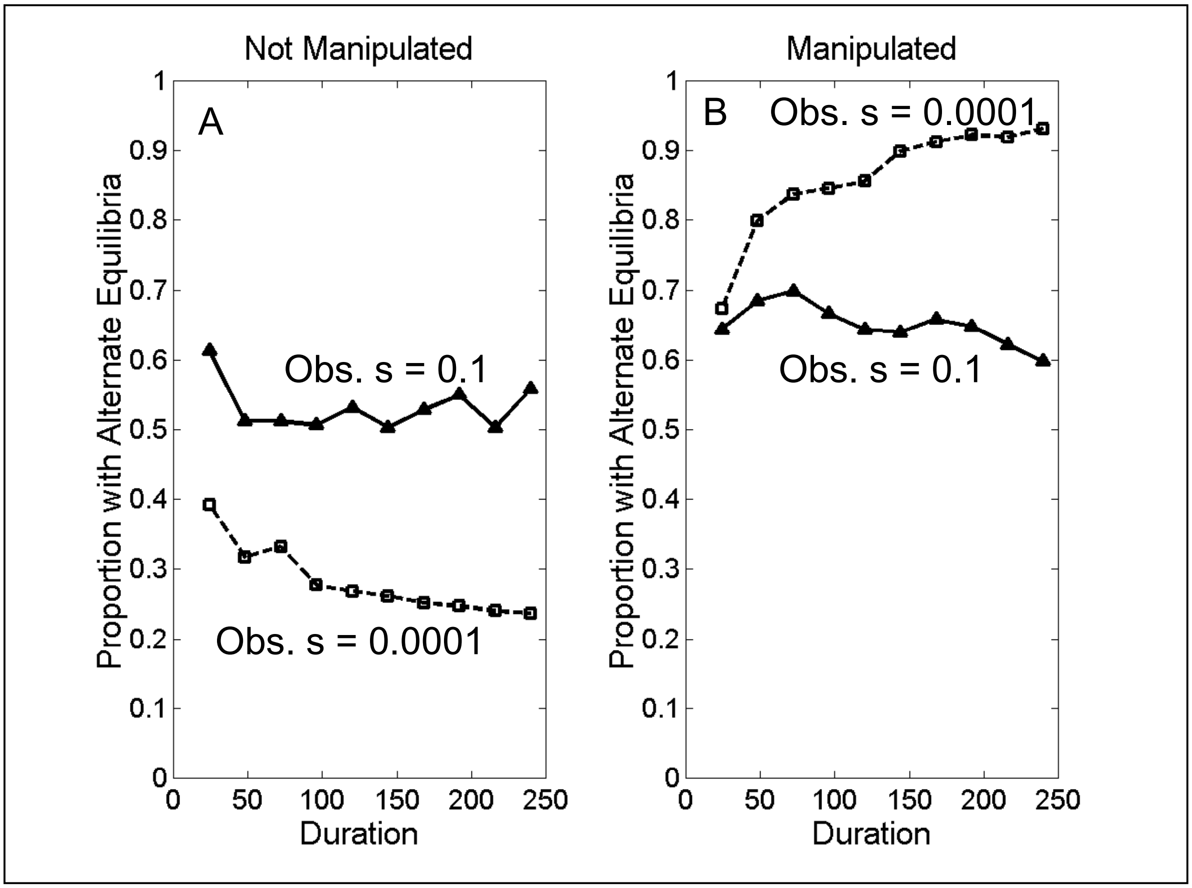

Figure 18. Indicators of recycling versus mass of phosphorus in the water of Lake Mendota during summer. A. Annual change in P mass (P_{t+1} - P_t). B. Net recycling (P_{t+1} - P_t - L_t + X_t). C. Residuals from the regression of (P_{t+1} + X_t) versus L_t. D. Residuals from the regression of (P_{t+1}- P_t + X_t) versus L_t. (Original)

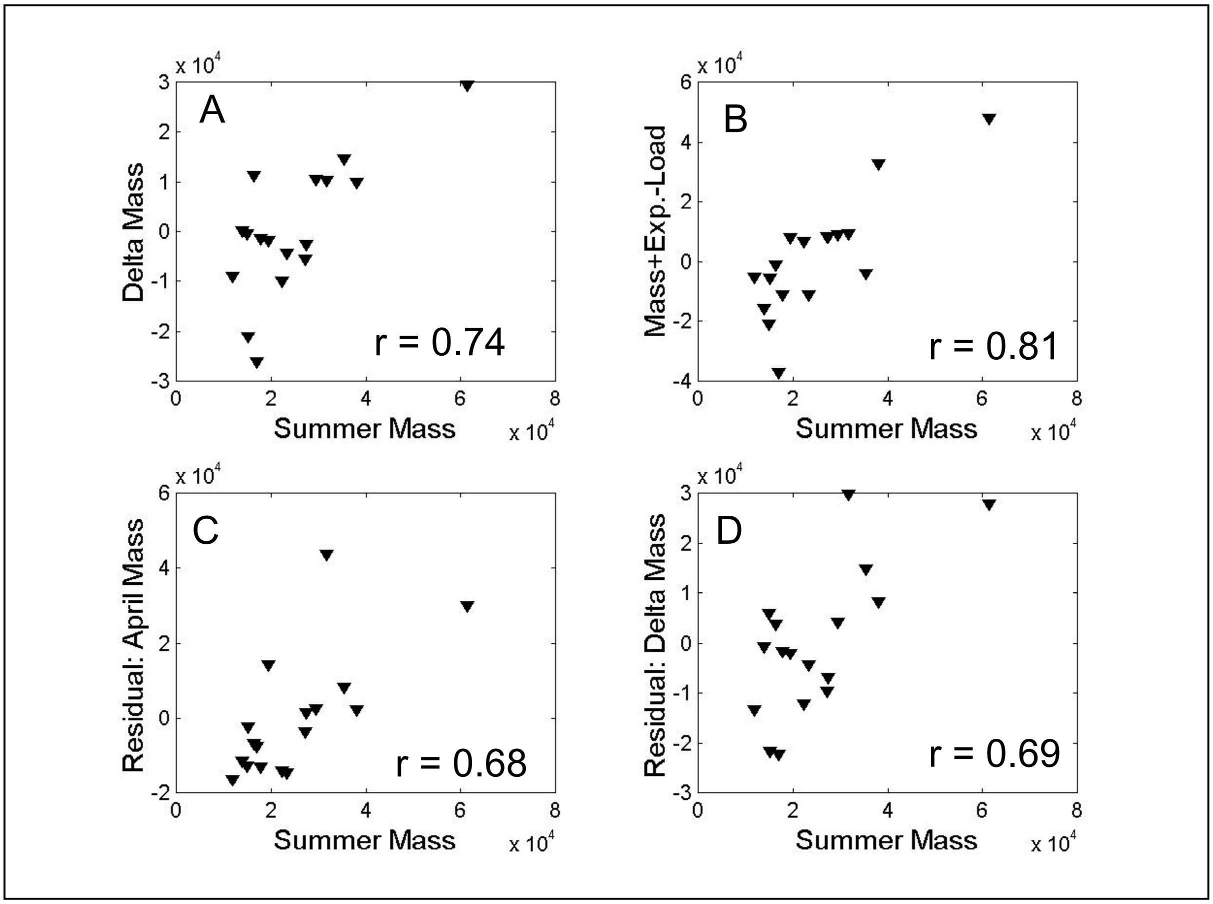

Figure 19. Observed P mass in Lake Mendota versus one-year-ahead prediction using the fitted equation 7 (triangles). Units of both axes are g m^{-2} y^{-1}. Line shows observations equal to predictions. (Original)

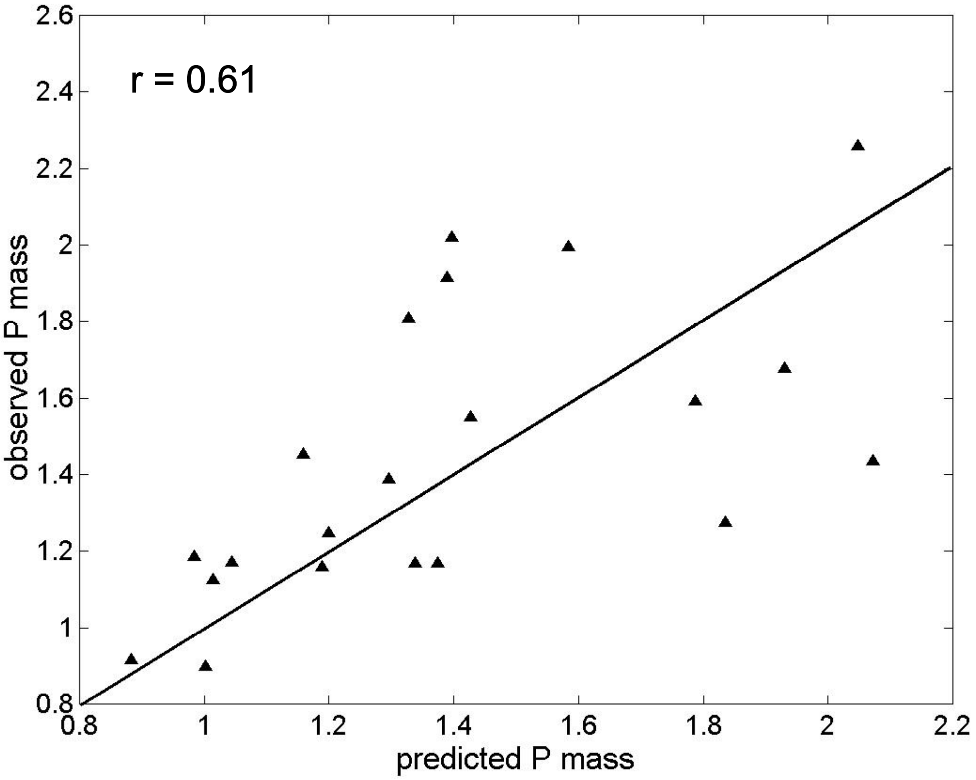

Figure 20. Probability distributions for the recycling parameter r from the 14 lakes studied by Nürnberg (1984) and the time series data for Lake Mendota. The distribution for the Lake Mendota data is a marginal distribution, integrated over the prior distributions of the other parameters. (Original)