A test for a regime shift

ANDREW R. SOLOW* AND ANDREW R. BEET
Marine Policy Center, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA

ABSTRACT
This note describes a formal statistical approach to testing for a regime shift based on population time series for a system of interacting species or groups. Under this approach, a regime shift is said to occur if the system switches abruptly from varying around one locally stable steady state to varying around another. This allows the problem to be formulated as testing for a changepoint in a vector autoregressive process. The test is illustrated using data from the North Sea covering the period 1963–97.

Key words: bootstrap, North Sea, regime shift, vector autoregressive process

INTRODUCTION
This note is concerned with the problem of detecting a regime shift in a collection of ecological time series. This problem has attracted considerable recent attention in the oceanographic literature (e.g. Francis et al., 1998; McGowan et al., 1998; Hare and Mantua, 2000; Reid et al., 2001). As Hare and Mantua (2000) pointed out, there is no common definition of a regime shift, nor is there a single method for detecting one. From a statistical point of view, the methods that have been proposed are ad hoc, in the sense that they are not based on a statistical model of the data. This poses a particular problem to the assessment of the significance of a possible regime shift identified by these methods. For example, Rudnick and Davis (2003) showed that the method proposed by Hare and Mantua (2000) for detecting a regime shift is sensitive to positive serial dependence in the time series.

The goal of this note is to propose a formal statistical approach to detecting a regime shift. Briefly, under this approach, a regime shift is defined as an abrupt shift in an ecosystem from one locally stable steady state to another locally stable steady state. To detect such a shift, we model the time series as a particular vector-valued stochastic process and test for a shift in the mean at an unknown time. The test is based on a combination of model fitting and a parametric bootstrap. We illustrate this approach using data from the North Sea ecosystem over the period 1963–97.

In related statistical work, Box and Tiao (1975) proposed the method that has come to be called intervention analysis for identifying a shift in the mean of a scalar-valued time series. Intervention analysis was extended to the case of vector-valued time series by Abraham (1980). The main difference between intervention analysis and the method proposed in this note is that, in intervention analysis, the time of the intervention (or regime shift) is assumed to be known, while one of the main goals of the method proposed here is to estimate the time of the regime shift.

METHOD
Consider a system of $k$ potentially interacting species. Let the random variables $X_{1t}, X_{2t}, \ldots, X_{kt}$ be the abundances (or log abundances) of these species in period $t$. A general stochastic model of the dynamics of this multi-species system is:

\[
X_{1t} = F_1(X_{1,t-1}, X_{2,t-1}, \ldots, X_{k,t-1}) + e_{1t}
\]
\[
X_{2t} = F_2(X_{1,t-1}, X_{2,t-1}, \ldots, X_{k,t-1}) + e_{2t}
\]
\[
\ldots
\]
\[
X_{kt} = F_k(X_{1,t-1}, X_{2,t-1}, \ldots, X_{k,t-1}) + e_{kt}
\]

where $F_1, F_2, \ldots, F_k$ are unknown non-linear functions and $e_{1t}, e_{2t}, \ldots, e_{kt}$ are normal process errors with mean 0 and unknown $k$-by-$k$ variance matrix $\Sigma$. This model can be compactly written in matrix notation as

\[
X_t = F(X_{t-1}) + e_t
\]

where $X_t = (X_{1t}, X_{2t}, \ldots, X_{kt})', F$ is now a vector-valued function, and $e_t$ is an $k$-variate normal process error with mean vector 0 and variance matrix $\Sigma$. We will assume the error process is serially independent. The extension to the case where this process is serially dependent is discussed briefly in the final section.
Let the vector $\mu$ satisfy:

$$\mu = F(\mu)$$  \hspace{1cm} (3)

That is, $\mu$ is a steady state of the deterministic system corresponding to the stochastic model (2). The dynamics of $X_t$ in the neighborhood of $\mu$ is approximated as

$$X_t \cong \mu + A(X_{t-1} - \mu) + \epsilon_t$$  \hspace{1cm} (4)

where the $k$-by-$k$ matrix $A = \left[ \frac{\partial F}{\partial X_{t-1} \mu} \right]$ is the (unknown) Jacobian of $F$. The model (4) is called a first-order vector autoregressive [VAR(1)] process (Reinsel, 1997). The process is stationary provided all of the eigenvalues of $A$ are less than 1 in modulus. This is the same condition for a univariate time series (Reinsel, 1997). The VAR(1) model can be fit by LS using the MATLAB package described in Schneider and Neumaier (2001). The LS estimates are asymptotically equivalent to the ML estimates and used in their place in forming $K$-statistic. Simulate a large number of time series from the null model for all possible values of $m$ and choose the fitted model with the largest overall likelihood. For technical reasons, this means that the standard distributional results for $\Lambda$ do not apply. Instead, the significance of the observed value of $\Lambda$ can be assessed through the following simulation procedure. Simulate a $k$-by-$k$ matrix $A$ from the model fitted under $H_0$. Fit the null and alternative models to the simulated data and find the value of $\Lambda$. Repeat the procedure a large number of times and estimate the significance level (or $P$-value) by the proportion of simulated series for which the value of $\Lambda$ is larger than that for the actual data. This is an example of a parametric bootstrap test (Efron and Tibshirani, 1994).

A limitation of this test is that to fit a unique least squares estimate of the $k$-by-$k$ matrix $A$ requires at least $k + 2$ observations. This is a minimal requirement and, needless to say, like all statistical methods, this one will be more powerful the longer the time series. As a result of this limitation, it is not practical to apply this test to a large number of time series as in Hare and Mantua (2000). The main use of the VAR(1) model is actually not in the estimation of $m$, but in the assessment of significance. To assess the significance of the results of any method for detecting a regime shift, it is necessary to account for...
dependence within and between the time series. The VAR(1) model provides a parsimonious description of this dependence.

AN ILLUSTRATION

As an illustration, the test outlined in the previous section was applied to \( k = 5 \) annual biological time series from the North Sea covering the \( N = 35 \) year period 1963–97. The time series used in this analysis were: (1) average phytoplankton color; and (2) total copepod abundance from the Continuous Plankton Recorder (CPR) and estimates of spawning stock biomass (SSB) for (3) cod (\textit{Gadus morhua}); (4) haddock (\textit{Melanogrammus aeglefinus}); and (5) herring (\textit{Clupea harengus}). This ordering of the variables will be maintained throughout the subsequent discussion. The first two time series are available at the website of the Sir Alister Hardy Foundation for Ocean Science (SAHFOS) and the last three are available at the UK website of the Department for Environment, Food, and Rural Affairs (DEFRA). These time series are shown in Fig. 1. We recognize that the selection of time series for inclusion in this kind of analysis is no small issue. However, as our primary goal here is illustrative, this selection, which was intended to cover a range of trophic levels, seems reasonable. Also, because the delay between spawning and recruitment is greater than 1 yr for cod and haddock, the VAR(1) model may not capture fully the

![Figure 1](image-url)

**Figure 1.** Observed values and fitted values under the null (solid) and alternative (dashed) hypotheses for (a) phytoplankton color; (b) total copepods; (c) cod; (d) haddock; and (e) herring. The fitted values for the two models are nearly indistinguishable. The vertical dotted line represents the year of the estimated changepoint. Also shown in (f) is the log likelihood for fitting the alternative model for different values of the changepoint.
dynamics of these species. We will return to these points in the final section.

The least squares estimate of the mean vector \( \mu \) under the null hypothesis of no regime shift is 
\[
(1.42, 269.4115, 9806.6)^T.
\]
The fitted values to the five time series under the null hypothesis are also shown in Fig. 1. The dominant eigenvalue of the fitted model has modulus 0.90, so the model is stationary, and the value of the maximized log likelihood is \(-243.1\). The value of the maximized log likelihood is shown in Fig. 1 for fitting the data under the alternative hypothesis of a regime shift for different values of the changepoint \( m \). The log likelihood has a maximum value of \(-218.3\), which occurs for \( m = 1987 \). The log likelihood for 1988 is only slightly lower. For an estimated changepoint in 1987, the least square estimates of the mean vectors \( \mu_1 \) and \( \mu_2 \) are 
\[
(1.32, 3154.9267, 0717.3)^T \quad \text{and} \quad (1.52, 275.7145, 6696.4)^T,
\]
respectively. In relative terms, the greatest difference between \( \mu_1 \) and \( \mu_2 \) are reductions of nearly 50% in mean cod and haddock SSB. The fitted values from this model are also shown in Fig. 1. The dominant eigenvalue of the first model has modulus 0.90 and the dominant eigenvalue of the second model has modulus 0.87.

The value of the LR statistic \( \Lambda \) for testing the null hypothesis of no regime shift is 49.7. Of 500 bootstrap samples simulated from the model fitted under the null model, 85 had values of \( \Lambda \) in excess of 49.7, for an estimated significance level of 0.17 with standard error of around 0.02. By conventional standards of significance, the data do not provide convincing evidence against the null hypothesis. Because it is based on different data and methods, this result is not directly comparable with those of previous studies. Despite these differences, Reid et al. (2001) selected 1988 as the most likely time for a regime shift in the North Sea. This is in good agreement with the results presented here. However, Reid et al. (2001) did not assess the significance of this change.

**DISCUSSION**

The purpose of this note has been to describe and illustrate a method for detecting a regime shift in population time series. Although this is certainly not the last word on this topic, we believe that the method described here has two distinct advantages over existing methods. First, it relies on the standard model of ecosystem dynamics to provide a statistical model on which a formal test can be based. Second, it relies on standard statistical theory to provide such a test. From a statistical perspective, two key features of the method are the allowance for serial dependence in the vector of population time series and the accounting for the selection of the most likely changepoint in the assessment of significance. In keeping with the literature in this area, the method proposed here is based on the definition of a regime shift as an abrupt change in steady state. As a practical matter, the performance of the method is unlikely to be disturbed by a transition period that is short in relation to the length of the observation period. However, if the definition is broadened to allow for a longer transition period, then it would be necessary to consider alternatives.

The basic method can be extended in a number of directions. Here, we mention four. First, the dynamics of one or more of the component series may involve lags greater than one period. For example, as noted, if there is a delay of \( \tau \) years between spawning and recruitment, then the SSB in year \( t \) will depend on SSB both in years \( t - 1 \) and \( t - \tau \). This can be accommodated by extending the order of the VAR model. Second, to the extent that the error process \( e_t \) reflects environmental forcing, it might be expected to exhibit positive serial dependence. The VAR(1) model can also be extended to accommodate this situation. For example, if \( e_t \) itself follows a VAR(1) model, then \( X_t \) is said to follow a vector autoregressive-moving average (VARMA) process of order 1 in the autoregressive part and order 1 in the moving average part (Reinsel, 1997). A VAR model of higher order can also be extended in this way. Third, the model can be modified to require the properties of the error process to be the same before and after the regime shift, so that \( \Sigma_1 = \Sigma_2 \). Fourth, the model can be extended to allow for the possibility of more than one regime shift. In general, allowing for multiple shifts requires some delicacy. For example, to avoid the presumably unrealistic case in which 2 regime shifts occur in close proximity, it may be necessary to impose a lower bound on permissible interval between the times of the shifts. Unfortunately, extensions of these kinds can introduce a large number of additional parameters, further limiting the applicability of the method to relatively short time series. Moreover, the great computational advantage of fitting VAR models by least squares is lost with VARMA models or even VAR models with special structure.

Finally, although a detailed discussion is beyond the scope of this note, a central issue in this kind of work is the selection of time series for inclusion in the analysis. One question is whether to include – as some studies have done – both biological and physical time series in the analysis. In our view, when interest centers on detecting an ecological regime shift, it is better to include only biological time series, although...
relating the results of the analysis to variations in the physical environment may be important. A second issue concerns the choice of biological time series. In this note, a regime shift has been defined as a shift in the steady state of an ecological system. The approach proposed here for detecting such a shift focuses directly on time series of abundance. As a shift in steady state abundance may be an indirect effect of a more direct ecological change (e.g. a shift in the stock–recruitment relationship of a fish stock), in some situations it may make sense to search for the direct effect itself. However, this presupposes some knowledge of the likely change.

ACKNOWLEDGEMENTS
The extremely helpful comments of Ian Perry and four anonymous reviewers are acknowledged with gratitude.

REFERENCES
Copyright of Fisheries Oceanography is the property of Blackwell Publishing Limited and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.