Stony Brook University

Novel nonparametric approaches to stock assessment and regime shift prediction

A Thesis Presented

by

Charles Thomas Perretti

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Master of Science

in

Marine and Atmospheric Science

Stony Brook University

December 2010

Stony Brook University

The Graduate School

Charles Thomas Perretti

We, the thesis committee for the above candidate for the

Master of Science degree, hereby recommend

acceptance of this thesis.

Dr. Stephan B. Munch, Thesis Advisor

Associate Professor

School of Marine and Atmospheric Science

Dr. Ellen K. Pikitch, Thesis Reader

Professor

School of Marine and Atmospheric Science

Dr. Perry deValpine

Assistant Professor

Environmental Science, Policy and Management

University of California, Berkeley

This thesis is accepted by the Graduate School

Lawrence Martin

Dean of the Graduate School

Abstract of the Thesis

Novel nonparametric approaches to stock assessment and regime shift prediction

by

Charles Thomas Perretti

Master of Science

in

Marine and Atmospheric Science

Stony Brook University

2010

Ecosystem dynamics are often complex, nonlinear, and characterized by critical thresholds or regime changes. Despite these difficulties, resource managers must accurately forecast species abundance and anticipate impending regime shifts in order to implement sustainable management plans.

In the first part of this thesis I explicitly describe a nonparametric method for multivariate forecasting which I call the MS-Map and evaluate its performance relative to a suite of parametric models. I found that, in the presence of noise, it is often possible to obtain more accurate forecasts from the MS-Map than from the model that was used to generate the data. The inclusion of additional species yielded a large improvement for the nonparametric MS-Map, a smaller improvement for the control model, and only a slight improvement for the alternative multi-species parametric model. When applied to rockfish larval abundance data from the CalCOFI survey, the performance of the MS-Map improved when additional species were included. These results suggest that flexible nonparametric modeling approaches should be considered for ecosystem management.

In the second part of this thesis, using the three-group fishery model previously studied by Biggs et al. (2009), I tested a suite of statistical regime shift indicators under the ecologically realistic conditions of high, correlated noise with short time series and a rapidly changing driving variable. I found that all indicators perform poorly under realistic conditions with the exception of the variance indicator. In contrast to expectations from previous work, the noise spectrum did not have a strong effect on indicator performance. The amount of data used to calculate the indicator had a large impact on performance. Also contrary to prior work, I found that the value of the spectral ratio was not a reliable indicator of an impending shift. Future research should focus on techniques that incorporate multiple data sources simultaneously, thus reducing the time needed to detect an impending shift.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would like to thank my family and girlfriend for their unwavering support. They are my foundation and inspiration, without which this would not have been possible.

I would like to thank my advisor and friend, Dr. Stephan Munch, for his willingness to accept me into his lab despite my unorthodox background. I benefited tremendously from his tireless instruction and, under his guidance, I learned how to be a scientist.

I would like to thank Dr. Ellen Pikitch for her mentorship and the many enriching experiences she afforded me through the Lenfest Forage Fish Task Force. The experience of working with a group of world-class scientists profoundly affected my career and scientific prespective. I would also like to thank Dr. Perry deValpine for his detailed and thoughtful comments which greatly expanded the scope and relevance of this work.

Lastly, I would like to thank everyone in the Munch lab for their assistance in the intellectual development and technical implementation of this thesis, and for greatly broadened my understanding of marine ecology and evolution.

I. Nonparametric multi-species forecasting evaluation

Introduction

An ecosystem-based approach to management requires forecasting the dynamics of all relevant species and the ability to anticipate the indirect effects of management decisions (Grumbine 1994, Slocombe 1998, Pikitch et al. 2004). Ecosystem dynamics are, however, often complex, nonlinear, and exhibit critical thresholds or regime changes, while our understanding of them is usually derived from short, noisy time series of just a few variables. Nevertheless, resource managers are charged with forecasting system dynamics in order to develop sustainable harvest regimes or conservation plans.

Many ecological modeling frameworks are available for forecasting, ranging from simple single-species models, to highly complex ecosystem models (Schaefer 1957, Christensen and Walters 2004, Fulton et al. 2005). Despite this range of tools, species collapses are not rare (Jackson et al. 2001, Myers and Worm 2003). Although collapses are sometimes attributed to dysfunctional governance (Safina and Klinger 2008), a critical factor leading to collapses is that the models used to make management decisions provided poor assessments or forecasts. One approach has been to move from single-species models to multi-species or ecosystem models in the hopes of increasing model realism. However, increasing model complexity will not necessarily lead to more accurate forecasts as network topology and the functional relationships between species are often highly uncertain (Ludwig et al. 1988). Moreover, apparently small changes in model structure may produce very large, qualitative changes in model predictions (Wood

and Thomas 1999). Nonparametric methods to forecasting (e.g. Härdle et al. 1997) are robust to this structural uncertainty and may provide a way forward.

Past studies have used nonparametric forecasting to identify chaos and, more generally, nonlinearities in ecological time series (Sugihara and May 1990, Sugihara et al. 1990), and although nonparametric techniques have gained wide popularity in other fields (Schreiber 1999), they have been largely ignored in natural resource management. One reason for this is that, as traditionally applied, nonparametric methods require long, highly precise time series, which are rare in ecology. Recent work has attempted to overcome this constraint by concatenating the time series of several ecologically similar species to form one long time series (Hsieh et al. 2008). Although early evaluations of this method have been promising, the process of combining time series from different species is potentially distorting, particularly if non-similar species are mistakenly included.

An alternate approach would be to use all of the data for all of the species in a given system by extending univariate embedding methods to multivariate time series. The idea of multivariate embedding has been used to improve forecasts in physics (Cao et al. 1998) and has been used for ecological time series, though without explanation (Dixon et al. 1999). To my knowledge, techniques for multivariate embedding have not as yet been described in an ecological context. Here, I explicitly describe a method for multivariate embedding which I will call the MS-Map and evaluate its performance relative to parametric models fit to real and simulated data.

Multivariate time delay embedding

Taken's theorem of state-space reconstruction (Takens 1981) allows for the reconstruction of a multivariate dynamical system through time delay embedding. As traditionally applied, a univariate time series is transformed into a set of time-delayed vectors: $X_t = [x_t, x_{t-\tau}, x_{t-2\tau},\ldots, x_{t-(E-1)\tau}]$, where x is a scalar time series, t is the time index, τ is the time lag, and E is the embedding dimension. Each vector X_t for $t=[1+(E-1)\tau,\ldots,N]$, where N is the length of the time series, is then embedded in an E dimensional phase-space, constructing an attractor that preserves the topological properties of the original system. Using this procedure, I generate an attractor for each species, then combine that information through a multivariate smoothing kernel I refer to as the MS-Map.

Predictions using the MS-Map are made by first dividing the time-delayed points for a given species into a training set and a test set. The length of the training set was varied from 45 points to 75 points, and the length of the test set was fixed at 60 points (see Table 1). The p-step prediction for each test point is the distance-weighted average of the positions of all training points p-steps forward in time. The weight of each train point is recalculated for each predictand and is given by:

$$w_{ij} = \exp\left[-\sum_{k} \frac{\theta_k d_{ijk}}{\bar{d}_{jk}}\right]$$

where θ_k describes the degree of local weighting for species k, d_{ijk} is the Euclidean distance between train point i and predictand j on the attractor reconstruction of species k, and $\bar{d}_{jk}$ is the average distance between predictand j and all other training points on the attractor reconstruction of species k. The weight for each training point is based on the relative distance of points from the predictand, thus it is a non-stationary smoothing kernel.

The MS-Map allows each species in the system to affect the weight parameter. A value of $\theta=0$ gives the mean of the training set as the prediction, while $\theta>0$ gives near-by points in phase space more weight, and is thus more nonlinear. If $\theta=0$ for a particular species then that species does not contribute to the forecast. The p-step predictions for each predictand are given by: $\hat{y}_j = \frac{1}{\sum_i w_{ij}} \sum_i w_{ij} x_{i+p}$, where w_{ij} is the weight of training point i for predictand j, and x_{i+p} is the p-step ahead value of training point i. A similar method was developed by Cao et al. (1998), and applied to several physical systems; however equal weights were given to all variables and the smoothing kernel was stationary.

To limit the computational requirements of the parameter search for this expository analysis, I restricted the attractor reconstructions to a time lag of unity and an embedding dimension of three. To avoid overfitting, the last fifteen points of the training set were excluded and used to determine the values of θ that resulted in the lowest forecast error. Those parameters were then used to forecast the out-of-sample test points. The time series of each species was scaled to [0,1] in order to standardize distances between points in phase space across species.

Parametric Models and Simulation Methods

I compared the forecast skill of the MS-Map model to three parametric models using simulated data from the three-species model by Hastings and Powell (1991) with harvesting, given by:

$$(1)\ \frac{dx}{dt} = x(1-x) - f_1(x)y - Fx$$

$$(2)\ \frac{dy}{dt} = f_1(x)y - f_2(y)z - d_1 y$$

$$(3)\ \frac{dz}{dt} = f_2(y)z - d_2 z$$

$$(4)\ f_i(u) = \frac{a_i u}{(1 + b_i u)}$$

The bottom species in the food chain (species x) is harvested at a constant rate (parameter F), which is analogous to a forage fish fishery. The model was numerically solved using a 4th-order runge-kutta method. One thousand parameter sets were drawn uniformly from the ranges given in Table 1. I focused on parameter sets that produced dynamics ranging from limit cycles to chaos and excluded any parameter sets that resulted in steady-state solutions, resulting in 317 unique parameter sets. I ran the model for an initial period of 50,000 time steps to avoid including transient dynamics. To model observation error, I multiplied the time series by log-normal noise for the range of standard deviations listed in Table 1, the highest intensity corresponded to a coefficient of variation of each species similar to that found in field studies (Reed and Hobbs, 2004).

I compared the forecast accuracy of the MS-Map algorithm to a single-species Schaefer model (Schaefer 1957), a three-species Lotka-Volterra model (LV), and the Hastings and Powell model (HP) that was used to generate the data. The functional form of the LV model is almost identical to the HP model, the only difference being that the Type II functional response of the HP model is replaced with a Type I functional response in the LV model. This level of similarity was chosen to assess the change in forecast accuracy due to a small difference in model structure. All parameters, including the initial conditions, were fit to the observed time series using maximum likelihood. Fitting nonlinear models to noisy data is notoriously difficult (Polansky et al. 2009), therefore I

initiated the fitting procedure at the correct initial conditions, and for the HP and LV models, the correct model parameters where ever possible in order to increase the probability that the global maximum of the likelihood function would be found. To compare performance across models I log-transformed the forecast root mean squared errors (RMSE) for each set of simulations and performed paired t-tests for all model pairs.

I evaluated the change in forecast accuracy of 3-step predictions (similar results were obtained using 1-step and 5-step predictions) when fitting the models with data from one, two and three species and for a range of training set lengths, noise intensities and harvest rates. I refer to training set lengths of 45, 60, and 75 points as short, medium and long time series, respectively. Low, medium, and high noise intensities refer to a standard deviation of the normal distribution of 0.05, 0.1, and 0.15. Harvest rates of 0.025, 0.05, and 0.075 correspond to low, medium, and high harvest rates, respectively.

CalCOFI Methods

The CalCOFI icthyoplankton survey is an ongoing 50+ year-long fishery-independent time series of larval abundance for the diverse group of fish that reside off the coast of southern California. I measured the change in forecast accuracy due to including additional species in the MS-Map model using larval abundance data from three rockfish species: shortbelly rockfish (*Sebastes jordani*), aurora rockfish, (*Sebastes aurora*), and bocaccio (*Sebastes bocaccio*). These species were selected from the 'expert opinion' database found in Hsieh et al. (2005) because they were observed in all sampled years and are viviparous, therefore they are a good proxy of adult abundance. Also, I

wanted to test the ability of the MS-Map to use species that were not predator and prey, thus extending its applicability. The time series for each species ranged from 1951-2007. The cruise abundance for each species is the average abundance in the 66 station area that has been in continual use since 1951. Annual averages for each species were calculated as the mean cruise abundance during the species' spawning months, as specified in Moser et al. 2001 (also see Hsieh et al. 2005). To standardize the phase-space distances across species I scaled each time series to [0,1]. The CalCOFI surveys were triennial from 1966 to 1984, therefore I interpolated the missing years using 1-d cubic Hermite splines (similar results were obtained using linear interpolation). The time series were initially divided into a 47 year training set and a 10 year test set. The last 15 years of the training set were used to find the optimal model parameters. One-step-ahead forecast accuracy was then evaluated using the out-of-sample test set.

Simulation Results

Overall results

In the simulations without noise the HP model generated perfect forecasts, even when fit with only one species time series. This was expected because the parameter search for the HP model was initiated at the correct values. The second best performing model was the MS-Map with an average error of 12%, 11%, and 9% when fit using one, two or three species, respectively. The distant third best was a tie between the Schafer model and the LV model fit to three species, both with an average error of 40%.

Surprisingly, for all simulations with noise, the MS-Map was more accurate than the HP model by, on average, 20% when fit using one species, 48% when using two

species, and by 38% when all three species were used; all differences were significant at the $p<0.0001$ level (Table 2). The mean percent error in forecasts for the one, two and three species MS-Map was 29%, 17%, and 15%, respectively, compared to the mean percent error for the HP model which was 36%, 33%, and 25%, respectively (Table 2). The difference was most dramatic under the high noise simulations where the MS-Map outperformed the HP model by 25%, 49%, and 44%, when fit with one, two or three species, respectively. Under the low noise simulations the difference in forecast accuracy was 8%, 40% and 16%, when fit with one, two or three species, respectively. All of the differences in forecast accuracy were statistically significant at the $p<0.0001$ level.

Overall, despite the similarity between the LV model and the HP model, the LV model failed to outperform even the single-species Schaefer model when using data from just one or two species. The forecast error for both models was always above 40% (Table 3). Only when the LV model was fit with three species did it outperform the Schaefer model, and then only by 6%. In comparing the Schaefer and the LV model, the better performing model was dependent on the amount of noise. Under the low noise simulations, the Schaefer model outperformed the LV model fit to one or two species ($p<0.001$) and was indistinguishable from the performance of the LV model fit to three species. In contrast, under the high noise simulations, the LV model outperformed the Schaefer model by approximately 10% for all species fits ($p<0.001$).

Effect of increasing the number of species used to fit the model

The MS-Map forecast accuracy improved significantly when additional species were used to fit the model. Across all simulations, there was a 40% improvement when the second species was added and an additional 10% improvement when the third species

was added ($p<0.0001$). The HP model also improved with each additional species, with an 8% improvement when a second species was added, and an additional 24% improvement when a third species was added ($p<0.0001$). The LV model did not show an improvement when going from a one to two species fit, however there was a 6% improvement when moving from one to three species ($p<0.0001$).

Effect of changing the time series length and fishing intensity

Over all simulations, the average forecast accuracy either increased slightly or did not change for all models when moving from the short time series (45 points) to the long time series (75 points); the increases were never more than 10% (Table 2). Interestingly, increasing the harvest rate resulted in a small improvement in forecast accuracy for the MS-Map model but the effect on the parametric models was mixed and not statistically significant.

CalCOFI Results

Overall, the three-species MS-Map exhibited a reduction in forecast error of 31% when compared to the one-species MS-Map (Figure 4). The improvement from moving to two species was slight, with only a 4% average improvement. The forecasts for the three-species model were more nonlinear than for the one-species model, with an average optimal weight parameter that was 60% larger.

Forecasts for each species differed when adding additional species (Figure 5). Forecasts for shortbelly rockfish were not improved by the addition of the other two species. In contrast, forecasts for aurora rockfish showed a large increase in accuracy when bocaccio was included to fit the model, however there was no increase in accuracy

when shortbelly rockfish was included. The aurora rockfish forecasts were improved when using three species, with nonlinear weighting being given to both the shortbelly rockfish and bocaccio time series. The bocaccio forecasts were improved by adding the shortbelly rockfish time series; however they were substantially worse when the aurora time series were added, possibly because the training set is too short to accurately estimate theta or because it only covers a small portion of the attractor. However, when all three time series were included there was an improvement similar to that of adding only shortbelly rockfish.

Discussion

These results show that, in the presence of noise, it is often possible to obtain more accurate forecasts from a nonparametric model than from the model that was used to generate the data. Moreover, the results show that a small change in the model structure can cause large declines in forecast accuracy. The LV model, which only differs from the HP model in the form of the functional response, had an average forecast error of over 40%, which was no better than the single-species Schaefer model. This indicates that a model which deviates even slightly from the correct structure may provide very poor forecasts, agreeing with past work (e.g. Wood and Thomas 1999, Skalski and Gilliam 2001). Given the high degree of uncertainty in the structure and dynamics of real ecosystems, I suggest that nonparametric methods and greater reliance on short-term forecasts may help improve resource management.

Considering the recent push towards ecosystem-based management, an important question is whether the forecast accuracy of models will improve by including the time

series of additional species. These results show that the inclusion of additional species yielded a large improvement for the nonparametric MS-Map, a smaller improvement for the HP model, and only a slight improvement for the LV model. For the MS-Map, additional species acted primarily as a noise reduction mechanism as the model was better able to determine which points in the state-space reconstruction of the target species should be weighted more heavily at each point in time. Since the observation errors were assumed to be independent, points which were spuriously close to the predictand due to noise contamination in the one-species map were less likely to be weighted heavily in the three species. This explains why the improvement was small when moving from the one species MS-Map to the three species MS-Map in the simulations without noise (<2% improvement), as opposed to the high noise simulations (greater than 40% improvement). For the HP model, additional species allowed for better parameter estimation as demonstrated by an average deviation in parameter fit of 19% in the high noise runs using one species, compared to 10% using three species. In contrast, there was only a small increase in the accuracy of the LV model with additional species as the model was rarely able to capture the dynamics of the system. Therefore, a flexible nonparametric framework may be the most appropriate way of integrating diverse data sources in an ecosystem management context, and may yield large improvements over highly complex ecosystem models.

When applied to time series of rockfish larval abundance, the mean error of the MS-Map when using three species was approximately 30% lower than when using one or two species. Importantly, although the effects of adding a second species were mixed, the inclusion of all three species always resulted in forecasts that were at least as accurate as

the single species MS-Map; i.e. unlike in the two-species case, there was no spurious fitting in the three species MS-Map. This suggests that the inclusion of more species may act as a buffer against the spurious weighting of non-relevant species. Interestingly, the improvement in forecast accuracy when adding species, and the magnitude of the weight parameter (θ), may be a useful way of identifying important ecological interactions and assessing interaction strengths.

Several future improvements can be made to the MS-Map. First, a spatially explicit version could provide large gains in forecast accuracy. For each species, both the location and abundance could be embedded in an n-dimensional space. Then, the MS-Map could be used analogously to forecast both location and abundance. Second, the MS-Map could easily be extended to include the magnitude and location of relevant physical variables such as seasonal temperature anomalies or the indices of the Pacific decadal oscillation.

In some situations, where either longer-term forecasts are needed or parts of a system are well approximated by parametric functions, it may be most appropriate to use semiparametric methods (e.g. Nelson et al. 2004) rather than nonparametric or fully parametric methods. Future research is needed to evaluate which management situations would most benefit from a nonparametric, vs. semiparametric, vs. fully parametric forecasting approach.

Lastly, if nonparametric methods are to gain widespread acceptance for resource management, new control rules based on attractor reconstruction must be proposed and evaluated. Nonparametric forecasting could be used to maximize the resilience of

ecosystems by eliminating harvest rules that would significantly alter the attractor shape or by avoiding high risk areas of the phase-space.

II. Phase shift indicators under ecologically realistic conditions

Introduction

Ecological phase shifts are characterized by sudden, long-lasting changes in ecosystem state. Examples of phase shifts include rapid transitions of kelp forests to barren substrate, of coral reefs to fleshy macroalgae, and of patchy shrubs to desert (Hughes 1994, Steneck et al. 2002, Konar and Estes 2003, Kefi et al. 2007). Recent analyses of fish populations and ocean circulation models indicate that bi-stability is likely to be the rule rather than the exception (Collie et al. 2004, Lenton et al. 2008). Moreover, many critical transitions are characterized by a phenomenon known as hysteresis, where in order to return the system to its original state the driving variable must be reverted well beyond the level which caused the shift. This irreversibility makes phase shifts potentially catastrophic for ecosystem users.

Forecasting phase shifts in nature is an exceptional challenge due to the lack of high resolution data and the inherent complexity of real ecosystems (deYoung et al. 2008). However, recent research has produced a suite of statistical indicators that show promise for the prediction of impending shifts (Scheffer et al. 2009). Many of the proposed indicators are based on the dynamical systems concept of 'critical slowing down' (Wissel 1984, Strogatz 1994) which describes the increase in the time it takes for a system to return to equilibrium after a small perturbation. This slowing down can be observed statistically in a time series as a rise in variance (Carpenter and Brock 2006), an increase in the AR1 coefficient (Dakos et al. 2008), or as a shift in skewness or kurtosis (Guttal and Jayaprakash 2008). The color of the power spectrum has also been proposed

as a leading indicator (Kleinen et al. 2002). As a system approaches a threshold the power at low frequencies is expected to increase faster than at high frequencies, therefore the ratio of the spectral density of low to high frequencies has been proposed as an indicator; the point at which this ratio exceeds one signaling an impending shift (Biggs et al. 2009). These indicators are ineffective for systems without a smooth potential, e.g. systems which exhibit global bifurcations (Graham and Tel 1984, Hastings and Wysham 2010). Nevertheless, the indicators are likely to be ecologically relevant given the numerous examples of real ecosystems that are well-approximated by models with smooth potentials (Scheffer 1998, Carpenter et al. 1999, Held and Kleinen 2004, van Nes and Scheffer 2005, Carpenter et al. 2008).

Previous evaluations of phase shift indicators have generally used long, high resolution time series of systems perturbed by weak, uncorrelated noise (Carpenter and Brock 2006, Guttal and Jayaprakash 2008, Biggs et al. 2009). Most natural populations, on the other hand, exhibit large temporal and spatial variability, are sampled at low frequencies (e.g. Sale et al. 1984, Gaines et al. 1985, Certain et al. 2007) and are subject to driving variables (e.g. harvesting) that change rapidly (Pope and Macer 1996). Moreover, the power spectral density of environmental time series, particularly marine time series, is strongly autocorrelated and well described by spectral exponents greater than zero (Steele 1985, Pelletier and Turcotte 1997, Cyr and Cyr 2003, Vasseur and Yodzis 2004). However, the efficacy of most phase shift indicators under ecologically realistic perturbations, noise levels, and noise spectra, is largely untested (but see Rudnick and Davis 2003).

Using the three-group fishery model previously studied by Biggs et al. (2009), I

tested a suite of indicators under ecologically realistic conditions. The system was simulated using historically reasonable rates of change in harvesting (Pope and Macer 1996) and typical levels of variability. I evaluated the performance of these indicators when using data sampled at a high frequency versus a realistic frequency and tested their sensitivity to autocorrelation in the process error. This represents an important step in determining the likely efficacy of these indicators for real ecosystems.

Methods

I evaluated a set of indicators for the multi-species fishery model derived by Carpenter and Brock (2004), and described in detail by Biggs et al. (2009). The system consists of three groups: adult piscivores (A), juvenile piscivores (J), and planktivores (F). The planktivores consume only juvenile piscivores while the adult piscivores consume both planktivores and juvenile piscivores. Planktivores and juvenile piscivores experience foraging arena dynamics in which they exchange between a refuge and a vulnerable state (Walters and Martell 2004). The system is represented as a coupled continuous-discrete model, where reproduction of piscivores occurs during the discrete interval while predation and fishing occur continuously. The dynamics of the continuous interval (known as the "monitoring interval") are given by:

$$\frac{dA}{dt} = -qEA$$

$$\frac{dF}{dt} = D_F(F_R - F) - c_{FA}FA + \sigma\frac{dW}{dt}$$

$$\frac{dJ}{dt} = c_{JA}JA - \frac{c_{JF}vFJ}{h + v + c_{JF}F}$$

Parameters are (q) catchability, (E) effort, (D_F) exchange rate between the refuge and foraging arena, (F_R) refuge reservoir, (c_{FA}) consumption rate of F by A, (c_{JA}) consumption rate of J by A, (c_{JF}) consumption rate of J by F, (v) rate that J become vulnerable to F, and (h) the rate that J enter their refuge. The additive noise ($\sigma \frac{dW}{dt}$) will take on the range of colors described below.

The discrete dynamics (known as the "maturation interval") are given by:

$$A_{t+1} = s(A_t + J_t)$$

$$F_{t+1} = F_t$$

$$J_{t+1} = fA_{t+1}$$

Parameters are (t) time interval, (s) survival of adult and juvenile piscivores between maturation intervals, and (f) fecundity of adult piscivores.

This model exhibits two distinct steady states, one dominated by piscivores, the other by planktivores. In the piscivore-dominant state, planktivore abundance is maintained at low levels by a large population of adult piscivores. In the planktivore–dominant state, piscivore abundance is suppressed through predation on the juvenile group. The transition between states is driven by fishing which targets only the adult piscivores.

I modified the original model by adding autocorrelated process noise to the planktivore dynamics. Noise in ecological time series is typically characterized by a power spectral density which scales with frequency (f) as $1/f^{\beta}$ (Inchausti and Halley 2002). Special cases include white ($\beta = 0$), pink ($\beta = 0.5$), red ($\beta = 1$) and brown ($\beta = 2$) noise. Estimates of β from environmental time series range from 0.25 to 1.6 (Steele 1985, Vasseur and Yodzis 2004), therefore, I tested the indicators with β ranging from 0

to 1.6 (Table 4).

In keeping with prior work (Biggs et al. 2009), I applied the process noise to the planktivore dynamics, using a symmetrically truncated normal distribution to prevent the ecological impossibility of negative population abundances at high noise levels while avoiding positive bias. The noise was generated using an autoregressive model of order one for which the variance was standardized across autocorrelation treatments. The indicators were evaluated over a range of noise intensities (see Tables 1 & 2), the maximum of which corresponds to a coefficient of variation for planktivores similar to that found in field studies (Reed and Hobbs, 2004).

The simulations were initiated near the equilibrium biomass values for qE=1.5, and run for 160 years after a burn-in period of 250 years. I induced the phase shift by steadily increasing qE each year after the burn-in. I chose an annual increase of 0.013 based on historical reconstructions of fishing mortality rates (Pope and Macer 1996). The "point of no return" is defined as the first year that a reduction of the harvest rate to 0.1 will not avert a phase shift. Beyond this point, the phase shift is effectively irreversible as no realistic reduction in harvest will allow the system to return to the piscivore-dominated state. In the deterministic model, this point is reached in year 58 of the simulation (qE=2.23). The "attractor switch point" occurs at year 23 of the simulation (qE=1.78), after which the relative abundances of piscivores and planktivores begin to flip as the system becomes dominated by planktivores (Figure 6).

I evaluated the probability that each indicator would provide sufficient warning to avert an impending phase shift. The indicators evaluated were (i) rise in variance, (ii) rise in the value of the return parameter of the AR1 model, (iii) shift in skewness and kurtosis,

and (iv) rise in the spectral density ratio. In following prior work (Biggs et al. 2009) I defined the spectral density ratio as the ratio of the spectral density at a frequency of 0.05 (low) to the spectral denity at a frequency of 0.5 (high). I expected to see a rise in variance and an increase in the return time as the system approached a critical threshold due to the progressive decline in system stability and weakening of attraction of the piscivore-dominant equilibrium. An increase in skewness and kurtosis was expected as the distribution of system states shifts in the direction of increasing planktivore abundance. The spectral density ratio was also expected to increase while the system approaches the critical threshold since low frequency directional movement begins to dominate the high frequency noise-induced fluctuations, although this was expected to be sensitive to the degree of noise autocorrelation.

The indicators were evaluated using both high and low amounts of data. The 'high data' simulations used 50 within-year data points each year (abundances during the "monitoring interval"); the 'low data' simulations used only one point per year (abundance at the maturation interval). In the high data simulations the indicators were calculated each year, while in the low data simulations the indicators were calculated using a moving window of fifteen years of annual data. Therefore in the low data simulations there were no indicator values before year fifteen.

I performed 1000 simulations for each noise, color, and data treatment. As an objective framework for evaluating each indicator, I tested them against a 'null model' in which no phase shift will occur (i.e. a model in steady state with qE = 1.5). Although past studies have reported that a spectral ratio above one is an indication of an impending shift, I found this was not true, as described below. Therefore, for all indicators I defined

a positive indication of a phase shift as an instance where the 10 year moving window of an indicator exhibited a slope significantly greater ($p < 0.1$) than expected under the null model; i.e. a positive indication of a phase shift occurred when the indicator was rising significantly faster than would be expected in a null model. Therefore, the first evaluation of an impending shift will occur at year 10 of the high data simulations and year 25 of the low data simulations (15 years to calculate the indicator, plus 10 years to calculate the slope in the low data simulations).

I also calculated the probability that a given indicator would produce a 'detection' significantly more frequently ($p < 0.1$) than the null model in the interval [0,T]. To do this, I (1) determined the distribution of the number of times the null model indicated a phase shift in [0,T]. Although ideally this would occur roughly 10% of the time, non-independence of successive indicator evaluations results in false detections 22% of the time, on average, up to the "point of no return" in the high noise simulations, (2) if an impending phase shift was concluded in any year in a simulation I set all subsequent years in that simulation to positive indications, (3) I calculated the probability that a given indicator would produce a 'detection' significantly more frequently than the null model in the interval [0,T] ($p=0.1$).

Results

In contrast to expectations based on prior work (Rudnick and Davis 2003) I found that the strength of the noise autocorrelation did not have a clear effect on the results of the indicator performance evaluations. Overall, the difference across autocorrelation treatments in the probability of concluding a phase shift was occurring up to the attractor

switch point and the "point of no return" was always less than 0.1. Therefore I will restrict further attention to results for β =1.6 for the remainder of the paper. Results for other noise colors were qualitatively the same.

Under the high-data, low-noise treatment the skewness indicator performed well, with a probability of detection greater than 0.82 at the attractor switch point and 0.84 at the "point of no return". Alternatively, the kurtosis indicator performed poorly, with a detection rate of 0.11 at the attractor switch point and the "point of no return", i.e. there were no subsequent indications of an impending phase shift after the attractor switch point (Figure 8a,c).

The skewness indicator deteriorated rapidly when subjected to increased noise, with a probability of detection at the "point of no return" of 0.25 and 0.28 in the medium and high noise treatments, respectively (Figure 8a). Interestingly, the kurtosis indicator improved, but was still weak in the medium and high noise treatments with a probability of detection of 0.29 prior to the "point of no return" in the medium and high noise simulations (Figure 8c).

Under the low-data, low-noise treatment the skewness and kurtosis indicators again performed poorly with a detection rate of less than 0.02 before the "point of no return". They were somewhat better at medium and high noise with a total detection rate prior to the "point of no return" of 0.17 and 0.15, respectively (Figure 8b,d).

Under the high-data, low-noise treatment the AR-1 indicator performed well, with a detection probability of 0.99 prior to the "point of no return". However, its performance was significantly weaker at medium and high noise, with a detection probability of 0.42 and 0.44 respectively (Figure 8e).

Under the low-data treatment, the AR1 indicator had a total detection rate prior to the "point of no return" of 0, 0.39, and 0.31 for the low, medium and high noise simulations, respectively (Figure 8f).

In agreement with prior analyses of this model (Biggs et al. 2009), the spectral density ratio performed well under all of the high-data treatments yielding detection probabilities of 0.8 or better (Figure 8g). However, under the low-data treatment the spectral ratio performs well only in the low noise scenario (detection probability ~0.99), under medium and high noise the detection probability drops to 0.31 and 0.29, respectively (Figure 8h).

Importantly, I found that a spectral ratio that exceeds one is not a reliable signal of an impending shift, as demonstrated in the equilibrium simulations (qE=1.69) shown in Figure 9 where the system does not exhibit a regime shift. Contrary to previous assertions (Biggs et al., 2009), it is possible for this model system to exhibit a spectral ratio above one for any sampling treatment or noise color, despite a zero probability of an impending shift. The value of the spectral ratio is determined by the value of parameter qE, not the future trajectory of the system.

Under both the high and low data treatments, a rise in variance was the best performing indicator overall with near certain detection prior to the "point of no return" across all noise treatments. Therefore, variance was the most reliable indicator of an impending phase shift and is the most robust to noise (Figure 8i,j).

Discussion

In our simulations, the performance of most indicators other than variance suffered greatly at realistic noise levels and often failed to provide sufficient warning to avert a phase shift. Increased noise had a significantly negative impact on the performance of the spectral ratio and AR1 indicators. In some cases additional noise improved the detection rate of the skewness and kurtosis indicators, possibly because the tails of the state distributions were not significantly affected in the low noise simulations and the skewness and kurtosis indicators are sensitive to such changes.

The sampling rate had a significant effect on indicator performance with an average decrease in detection probability of 0.4 at the "point of no return" when moving from high to low data treatments. This was particularly true for the spectral ratio, where the detection probability at low noise and high data was 72% lower than at low noise and low data (Figure 8g,h). Encouragingly, I found that the variance indicator was robust to noise with a detection probability prior to the "point of no return" exceeding 0.9 in all treatments.

Overall, the strength of the noise autocorrelation does not significantly affect the performance of the indicators. This is particularly surprising for the spectral ratio which was expected to depend heavily on noise color. This appears to be because, even in the highest autocorrelation treatment, the color of the planktivore time series at low abundance is determined almost completely by the dynamics of the system, i.e. the color of the noise does not significantly affect the color of the planktivore time series as any increase in planktivore abundance is quickly offset by an increase in predation by adult piscivores. Therefore, at low planktivore abundance, the spectrum of the time series is determined by the proximity to the critical threshold, which is itself controlled by the

harvest rate on piscivores, not the noise spectrum. This effect is muted after the phase shift because piscivores no longer assert a strong effect on planktivore dynamics, therefore planktivores are more heavily affected by the noise color. Thus, although the degree of noise autocorrelation does not have a significant effect on indicator performance in our system, it may be important in other systems where the spectral density of the species time series is determined by the noise autocorrelation, e.g. populations that are strongly influenced by environmental variation rather than predation.

A persistent problem in phase shift detection is the inability of indicators to provide an *a priori* critical value for management action. It was previously suggested that the spectral ratio could overcome this problem by signaling a shift when the ratio exceeded one. However, the value of the ratio alone is not a reliable indicator of an impending shift as it can take on values above one at equilibrium. Therefore, I was left with only the ability to look at trends in indicators over time. One solution would be to use the change in the indicator relative to its range of values over a putative equilibrium period as an indication of an impending shift. However, this comparison could only indicate the probability that the system has changed from the equilibrium period, not the probability of an impending shift. Therefore, a necessary next step in phase shift prediction is the development of indicators that provide a definite signal of an upcoming shift.

I suggest that future work focus on the use of multivariate time series and state-space reconstruction techniques like those outlined in Hsieh et al. (2008) to uncover the global dynamics of the system using data from a subset of the active variables. None of the indicators tested here explicitly attempt to incorporate data from multiple species

simultaneously which is likely to improve detection rates by providing additional information about the stability of the system. One approach would be to use a measure of the change in the effective dimension of the system as an indication of an impending shift. In the system studied here, as the critical threshold is approached, the importance of piscivores for the regulation of planktivores decreases and the system shifts from two to three dimensions down to one. Such a method may be able to detect non-local bifurcations which cannot be detected using the indicators tested here (see Hastings and Wysham 2010). In addition, spatial state-space reconstruction techniques can be used for systems that lack sufficiently long time series. For example, comparisons of the number of active dimensions for similar systems along a gradient of habitat degradation (e.g. Sandin et al. 2008) may be able to provide an instantaneous measure of phase shift proximity.

In summary, I found that, overall, the sampling rate and the noise intensity had large negative impacts on indicator performance. The variance indicator was the most robust to noise and changes in sampling rate and performed well across all treatments. As real systems are not accompanied by a null model, future work should focus on the identification of critical values which resource managers can use as definite signals of an upcoming shift. Also, considering recent advances in remote sensing and ecological informatics, future work should focus on developing state-space reconstruction techniques that can utilize multiple data sources simultaneously, thus reducing the time needed to detect an impending shift.

References

Biggs, R., Carpenter, S.R. & Brock, W.A. (2009). Turning back from the brink: Detecting an impending regime shift in time to avert it. *Proceedings of the National Academy of Sciences*, 106, 826 -831.

Cao, L., Mees, A. & Judd, K. (1998). Dynamics from multivariate time series. *Physica D: Nonlinear Phenomena*, 121, 75-88.

Carpenter, S.R & Brock, W.A. (2004). *Spatial complexity, resilience and policy diversity : fishing on lake-rich landscapes*. Wisconsin Madison - Social Systems.

Carpenter, S.R. & Brock, W.A. (2006). Rising variance: a leading indicator of ecological transition. *Ecol. Lett*, 9, 311-318.

Carpenter, S.R., Brock, W.A., Cole, J.J., Kitchell, J.F. & Pace, M.L. (2008). Leading indicators of trophic cascades. *Ecology Letters*, 11, 128-138.

Carpenter, S.R., Ludwig, D. & Brock, W.A. (1999). Management of eutrophication for lakes subject to potentially irreversible change. *Ecological Applications*, 9, 751-771.

Certain, G., Bellier, E., Planque, B. & Bretagnolle, V. (2007). Characterising the temporal variability of the spatial distribution of animals: an application to seabirds at sea. *Ecography*, 30, 695-708.

Christensen, V. & Walters, C.J. (2004). Ecopath with Ecosim: methods, capabilities and limitations. *Ecological Modelling*, 172, 109-139.

Collie, J.S., Richardson, K. & Steele, J.H. (2004). Regime shifts: Can ecological theory illuminate the mechanisms? *Progress In Oceanography*, 60, 281-302.

Cyr, H. & Cyr, I. (2003). Temporal scaling of temperature variability from land to oceans.

Evolutionary Ecology Research, 5, 1183-1197.

Dakos, V., Scheffer, M., van Nes, E.H., Brovkin, V., Petoukhov, V. & Held, H. (2008). Slowing down as an early warning signal for abrupt climate change. *Proceedings of the National Academy of Sciences*, 105, 14308-14312.

deYoung, B., Barange, M., Beaugrand, G., Harris, R., Perry, R.I., Scheffer, M. et al. (2008). Regime shifts in marine ecosystems: detection, prediction and management. *Trends in Ecology & Evolution*, 23, 402-409.

Dixon, P.A., Milicich, M.J. & Sugihara, G. (1999). Episodic Fluctuations in Larval Supply. *Science*, 283, 1528-1530.

Fulton, E.A., Smith, A.D. & Punt, A.E. (2005). Which ecological indicators can robustly detect effects of fishing? *ICES J. Mar. Sci.*, 62, 540-551.

Gaines, S., Brown, S. & Roughgarden, J. (1985). Spatial variation in larval concentrations as a cause of spatial variation in settlement for the barnacle, Balanus glandula. *Oecologia*, 67, 267-272.

Graham, R. & Tel, T. (1984). Existence of a Potential for Dissipative Dynamical Systems. *Phys. Rev. Lett.*, 52, 9.

Grumbine, R.E. (1994). What Is Ecosystem Management? *Conservation Biology*, 8, 27-38.

Guttal, V. & Jayaprakash, C. (2008). Changing skewness: an early warning signal of regime shifts in ecosystems. *Ecology Letters*, 11, 450-460.

H. G. Moser, R. L. Charter, P. E. Smith, D.A. Ambrose, W. Watson, S. R. Charter et al. (n.d.). Distributional atlas of fish larvae and eggs in the Southern California Bight region: 1951-1998.

Härdle, W., Lütkepohl, H. & Chen, R. (1997). A Review of Nonparametric Time Series Analysis. *International Statistical Review / Revue Internationale de Statistique*, 65, 49-72.

Hastings, A. & Powell, T. (1991). Chaos in a Three-Species Food Chain. *Ecology*, 72, 896-903.

Hastings, A. & Wysham, D.B. (2010). Regime shifts in ecological systems can occur with no warning. *Ecology Letters*, 13, 464-472.

Held, H. & Kleinen, T. (2004). Detection of climate system bifurcations by degenerate fingerprinting. *Geophys. Res. Lett.*, 31, 4 PP.

Hsieh, C., Reiss, C., Watson, W., Allen, M.J., Hunter, J.R., Lea, R.N. et al. (2005). A comparison of long-term trends and variability in populations of larvae of exploited and unexploited fishes in the Southern California region: A community approach. *Progress in Oceanography*, 67, 160-185.

Hsieh, C., Anderson, C. & Sugihara, G. (2008). Extending nonlinear analysis to short ecological time series. *Am. Nat*, 171, 71-80.

Hughes, T.P. (1994). Catastrophes, Phase Shifts, and Large-Scale Degradation of a Caribbean Coral Reef. *Science*, 265, 1547-1551.

Inchausti, P. & Halley, J. (2002). The long-term temporal variability and spectral colour of animal populations. *Evolutionary Ecology Research*, 4, 1033-1048.

Jackson, J.B.C., Kirby, M.X., Berger, W.H., Bjorndal, K.A., Botsford, L.W., Bourque, B.J. et al. (2001). Historical Overfishing and the Recent Collapse of Coastal Ecosystems. *Science*, 293, 629-637.

Kefi, S., Rietkerk, M., Alados, C.L., Pueyo, Y., Papanastasis, V.P., ElAich, A. et al.

(2007). Spatial vegetation patterns and imminent desertification in Mediterranean arid ecosystems. *Nature*, 449, 213-217.

Kleinen, T., Held, H. & Petschel-Held, G. (2003). The potential role of spectral properties in detecting thresholds in the Earth system: application to the thermohaline circulation. *Ocean Dynamics*, 53, 53-63.

Konar, B. & Estes, J.A. (2003). The Stability of Boundary Regions Between Kelp Beds and Deforested Areas. *Ecology*, 84, 174-185.

Ludwig, D., Walters, C.J. & Cooke, J. (1988). Comparison of two models and two estimation methods for catch and effort data. *Natural Resource Modeling*, 2, 457-498.

Myers, R.A. & Worm, B. (2003a). Rapid worldwide depletion of predatory fish communities. *Nature*, 423, 280-283.

Myers, R.A. & Worm, B. (2003b). Rapid worldwide depletion of predatory fish communities. *Nature*, 423, 280-283.

Nelson, W.A., McCauley, E. & Wimbert, J. (2004). Capturing dynamics with the correct rates: inverse problems using semiparametric approaches. *Ecology*, 85, 889-903.

Pelletier, J.D. & Turcotte, D.L. (1997). Long-range persistence in climatological and hydrological time series: analysis, modeling and application to drought hazard assessment. *Journal of Hydrology*, 203, 198-208.

Pikitch, E.K., Santora, C., Babcock, E.A., Bakun, A., Bonfil, R., Conover, D.O. et al. (2004). Ecosystem-Based Fishery Management. *Science*, 305, 346-347.

Polansky, L., de Valpine, P., Lloyd-Smith, J.O. & Getz, W.M. (2009). Likelihood ridges and multimodality in population growth rate models. *Ecology*, 90, 2313-2320.

Pope, J.G. & Macer, C.T. (1996). An evaluation of the stock structure of North Sea cod, haddock, and whiting since 1920, together with a consideration of the impacts of fisheries and predation effects on their biomass and recruitment. *ICES J. Mar. Sci.*, 53, 1157.

Porporato, A. & Ridolfi, L. (2001). Multivariate nonlinear prediction of river flows. *Journal of Hydrology*, 248, 109-122.

Reed, D.H. & Hobbs, G.R. (2004). The Relationship Between Population Size and Temporal Variability in Population Size. *Animal Conservation*, 7, 1-8.

Rudnick, D.L. & Davis, R.E. (2003). Red noise and regime shifts. *Deep Sea Research Part I: Oceanographic Research Papers*, 50, 691-699.

Safina, C. & Klinger, D.H. (2008). Collapse of bluefin tuna in the Western Atlantic. *Conserv. Biol*, 22, 243-246.

Sale, P.F., Doherty, P.J., Eckert, G.J., Douglas, W.A. & Ferrell, D.J. (1984). Large scale spatial and temporal variation in recruitment to fish populations on coral reefs. *Oecologia*, 64, 191-198.

Sandin, S.A., Smith, J.E., DeMartini, E.E., Dinsdale, E.A., Donner, S.D., Friedlander, A.M. et al. (2008). Baselines and Degradation of Coral Reefs in the Northern Line Islands. PLoS ONE, 3, e1548.

Schaefer, M. (1957). Some Considerations of Population Dynamics and Economics in Relation to the Management of the Commercial Marine Fisheries. *Journal of the Fisheries Research Board of Canada*, XIV, 669-681.

Scheffer, M. (1998). *Ecology of shallow lakes*. Springer.

Scheffer, M., Carpenter, S., Foley, J.A., Folke, C. & Walker, B. (2001). Catastrophic

shifts in ecosystems. *Nature*, 413, 591-596.

Schreiber, T. (1999). Interdisciplinary application of nonlinear time series methods. *Physics Reports*, 308, 1-64.

Skalski, G.T. & Gilliam, J.F. (2001). Functional responses with predator interference: viable alternatives to the holling type II model. *Ecology*, 82, 3083-3092.

Slocombe, D.S. (1998). Defining Goals and Criteria for Ecosystem-Based Management. *Environmental Management*, 22, 483-493.

Steele, J.H. (1985). A comparison of terrestrial and marine ecological systems. *Nature*, 313, 355-358.

Steneck, R.S., Graham, M.H., Bourque, B.J., Corbett, D., Erlandson, J.M., Estes, J.A. et al. (2002). Kelp Forest Ecosystems: Biodiversity, Stability, Resilience and Future. *Environmental Conservation*, 29, 436-459.

Strogatz, S.H. (2001). *Nonlinear Dynamics And Chaos: With Applications To Physics, Biology, Chemistry, And Engineering*. 1st edn. Westview Press.

Sugihara, G., Grenfell, B. & May, R.M. (1990). Distinguishing error from chaos in ecological time series. *Philos. Trans. R. Soc. Lond., B, Biol. Sci*, 330, 235-251.

Sugihara, G. & May, R.M. (1990). Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. *Nature*, 344, 734-741.

Takens, F. (1981). Detecting strange attractors in turbulence. *Dynamical Systems and Turbulence*, 366-381.

van Nes, E.H. & Scheffer, M. (2005). Implications of spatial heterogeneity for catastrophic regime shifts in ecosystems. *Ecology*, 86, 1797-1807.

Vasseur, D.A. & Yodzis, P. (2004). The color of environmental noise. *Ecology*, 85, 1146-

1152.

Walters, C.J. & Martell, S.J.D. (2004). *Fisheries ecology and management*. Princeton University Press.

Wissel, C. (1984). A universal law of the characteristic return time near thresholds. *Oecologia*, 65, 101-107.

Wood, S.N. & Thomas, M.B. (1999). Super–sensitivity to structure in biological models. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 266, 565-570.

Table 1 Parameter ranges for the Hastings and Powell (1991) model

Parameter	Values
a_1	[3.5 to 6.5]
a_2	[0.7 to 0.13]
b_1	[2.1 to 3.9]
b_2	[1.4 to 2.6]
d_1	[0.28 to 0.52]
d_2	[0.007 to 0.013]
F	[0.025 to 0.075]
σ	[0.05 to 0.15]
θ	[0 to 500]

Table 2 Mean RMSE for all models, fit with one, two, and three species, denoted as (1), (2) and (3). LV is the Lotka-Volterra model, HP is the Hastings-Powell model.

Forecasting method	Overall	No noise	Low noise	High noise	Short series	Long series	Low harvest	High harvest
MS-Map (1)	0.21	0.09	0.17	0.24	0.22	0.20	0.21	0.20
MS-Map (2)	0.12	0.08	0.10	0.16	0.15	0.11	0.14	0.11
MS-Map (3)	0.11	0.07	0.08	0.15	0.13	0.10	0.12	0.10
Schaefer (1)	0.32	0.29	0.30	0.35	0.32	0.32	0.32	0.32
LV (1)	0.32	0.36	0.34	0.32	0.33	0.32	0.32	0.34
LV (2)	0.32	0.35	0.34	0.32	0.32	0.32	0.31	0.34
LV (3)	0.30	0.29	0.29	0.32	0.30	0.31	0.30	0.32
HP (1)	0.26	0.00	0.18	0.32	0.26	0.25	0.25	0.26
HP (2)	0.24	0.00	0.16	0.30	0.24	0.24	0.23	0.23
HP (3)	0.18	0.00	0.10	0.26	0.17	0.18	0.18	0.17

Table 3 Mean percent error for all models, fit with one, two and three species, denoted as (1), (2) and (3). LV is the Lotka-Volterra model, HP is the Hastings-Powell model.

Forecasting method	Overall	No noise	Low noise	High noise	Short series	Long series	Low harvest	High harvest
MS-Map (1)	29.2%	12.2%	23.6%	33.3%	30.6%	27.8%	29.2%	27.8%
MS-Map (2)	16.7%	10.7%	13.9%	22.2%	20.8%	15.3%	19.4%	15.3%
MS-Map (3)	15.3%	9.1%	11.1%	20.8%	18.1%	13.9%	16.7%	13.9%
Schaefer (1)	44.4%	39.8%	41.7%	48.6%	44.4%	44.4%	44.4%	44.4%
LV (1)	44.4%	49.3%	47.2%	44.4%	45.8%	44.4%	44.4%	47.2%
LV (2)	44.4%	49.0%	47.2%	44.4%	44.4%	44.4%	43.1%	47.2%
LV (3)	41.7%	39.9%	40.3%	44.4%	41.7%	43.1%	41.7%	44.4%
HP (1)	36.1%	0.0%	25.0%	44.4%	36.1%	34.7%	34.7%	36.1%
HP (2)	33.3%	0.0%	22.2%	41.7%	33.3%	33.3%	31.9%	31.9%
HP (3)	25.0%	0.0%	13.9%	36.1%	23.6%	25.0%	25.0%	23.6%

Table 4 Probability of detecting a shift for the *high* data simulations for the period up to one year before the point of no return.

Indicator	Low Noise ($\sigma = 0.1$)	Medium Noise ($\sigma = 2.2$)	High Noise ($\sigma = 4.3$)
Skew	0.84	0.28	0.26
Kurtosis	0.11	0.29	0.29
AR-1	1.00	0.47	0.44
Spectral Ratio	1.00	0.98	0.80
Variance	1.00	1.00	1.00

Table 5 Probability of detecting a shift for the *low* data simulations for the period up to one year before the point of no return.

Indicator	Low Noise ($\sigma = 0.1$)	Medium Noise ($\sigma = 2.2$)	High Noise ($\sigma = 4.3$)
Skew	0.01	0.17	0.23
Kurtosis	0.01	0.15	0.22
AR-1	0.00	0.39	0.31
Spectral Ratio	0.99	0.31	0.29
Variance	1.00	0.97	0.94

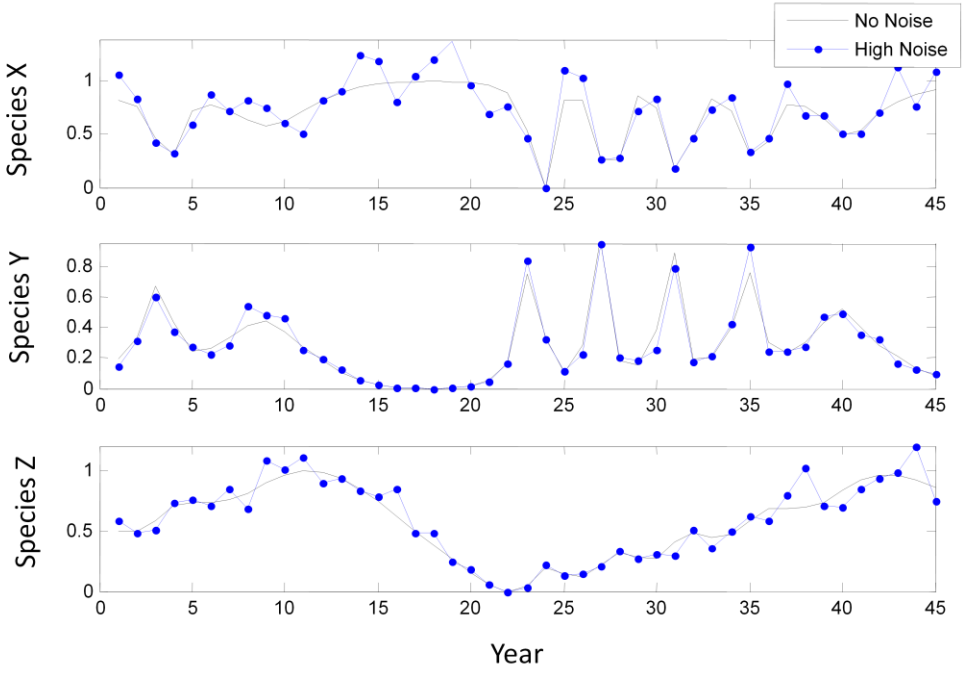

Figure 1 A typical time series for the short (45 year) simulations. High noise time series is in blue, deterministic time series is in black.

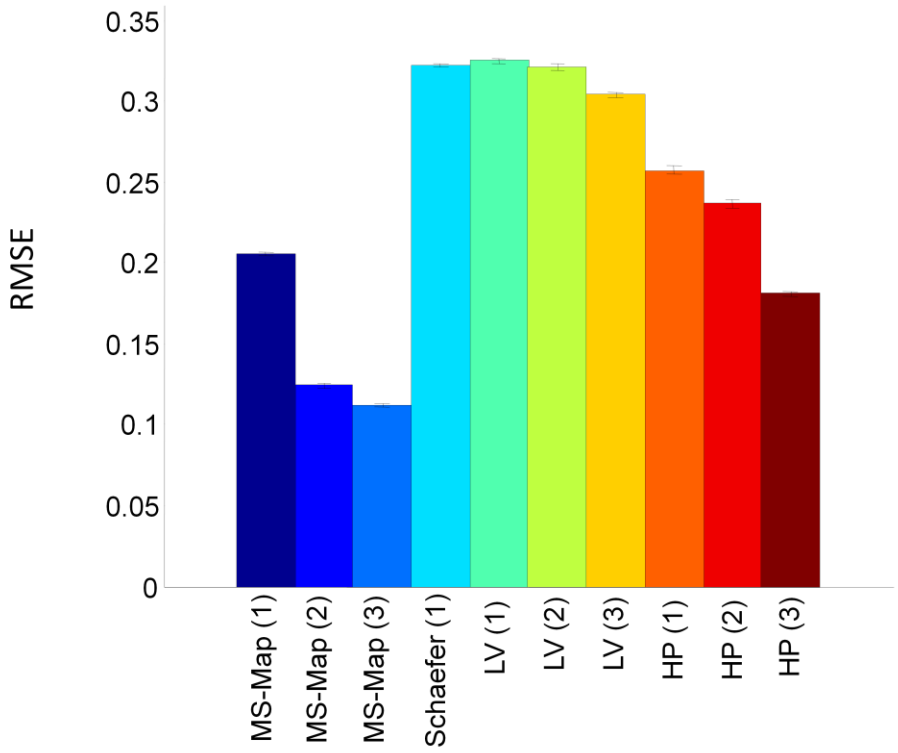

Figure 2 Mean RMSE for all simulations with noise for the HP system. Over all simulations with noise the MS-Map outperformed all other models. For simulations without noise (not shown) the HP model generated perfect forecasts.

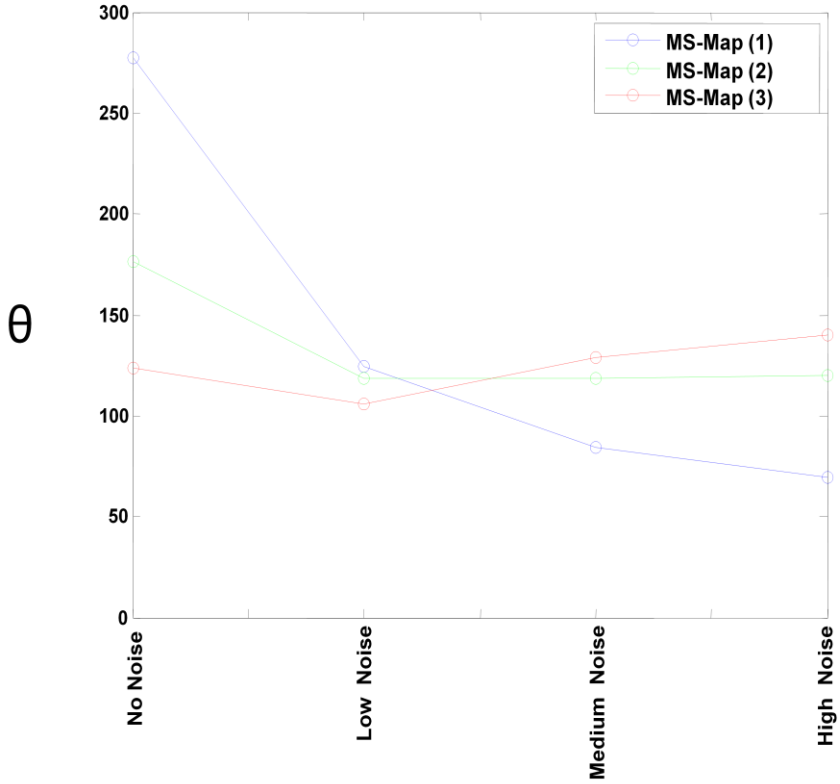

Figure 3 Change in the weight parameter (θ) of the MS-Map vs. noise level. A large θ generates forecasts that give points close to the predictand more weight, i.e. forecasts that are more nonlinear. The three-species MS-Map was able to maintain the nonlinearity of the forecasts through the high noise simulations.

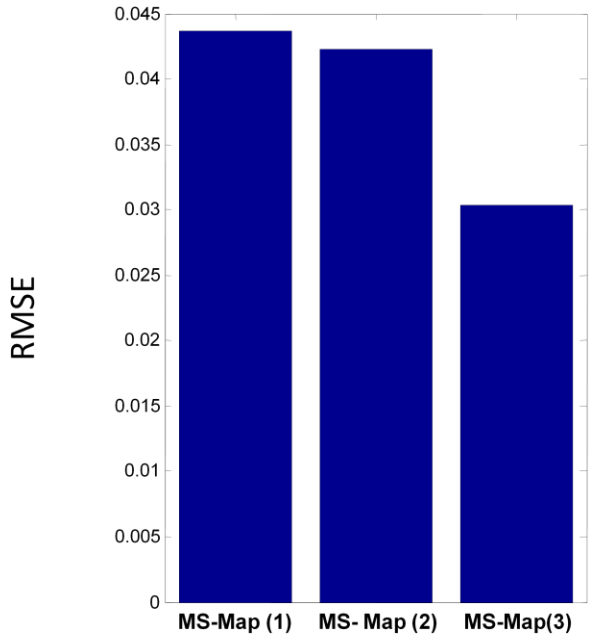

Figure 4 Mean RMSE for all combinations of the CalCOFI rockfish larval abundance data. The one-species, two-species and three-species MS-Maps RMSEs are denoted as (1), (2) and (3).

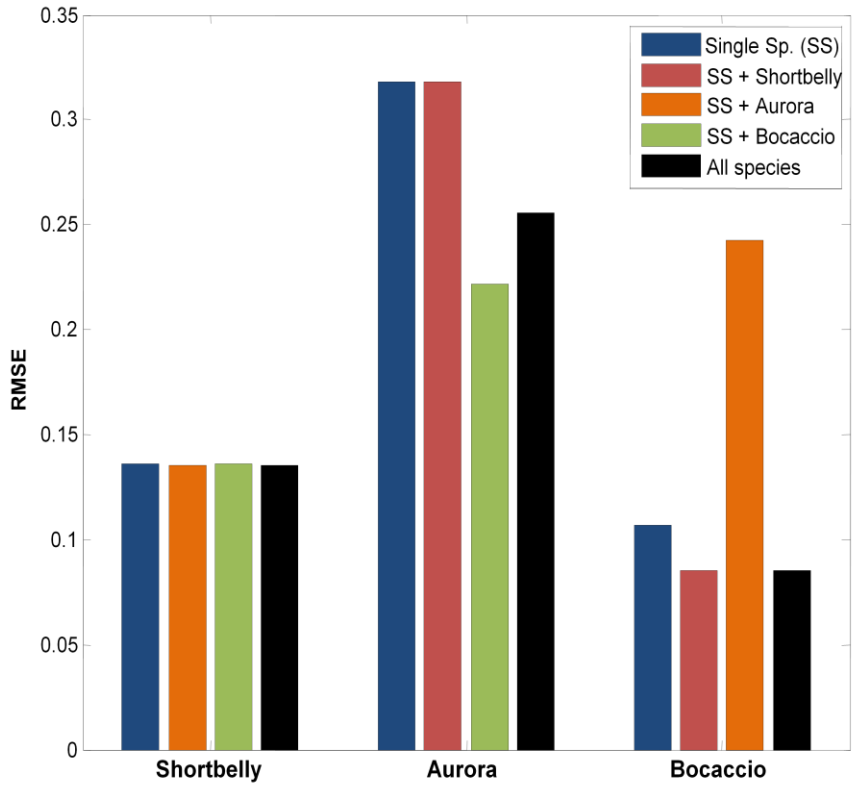

Figure 5 RMSE of the MS-Map forecasts for each combination of the three rockfish species. The x-axis gives the species being forecasted. Each group of bars describes the forecast accuracy when using just the species being forecasted (blue bars), using the species being forecasted and one other species (red, orange, and green bars), and using all three species (black bars). Shortbelly forecasts did not improve when adding species. Aurora forecasts did not improve when including Shortbelly, but did improve when adding Bocaccio and all three species. Boccacio forecasts were not improved by adding Aurora, but did improve when adding shortbelly and when including all three species.

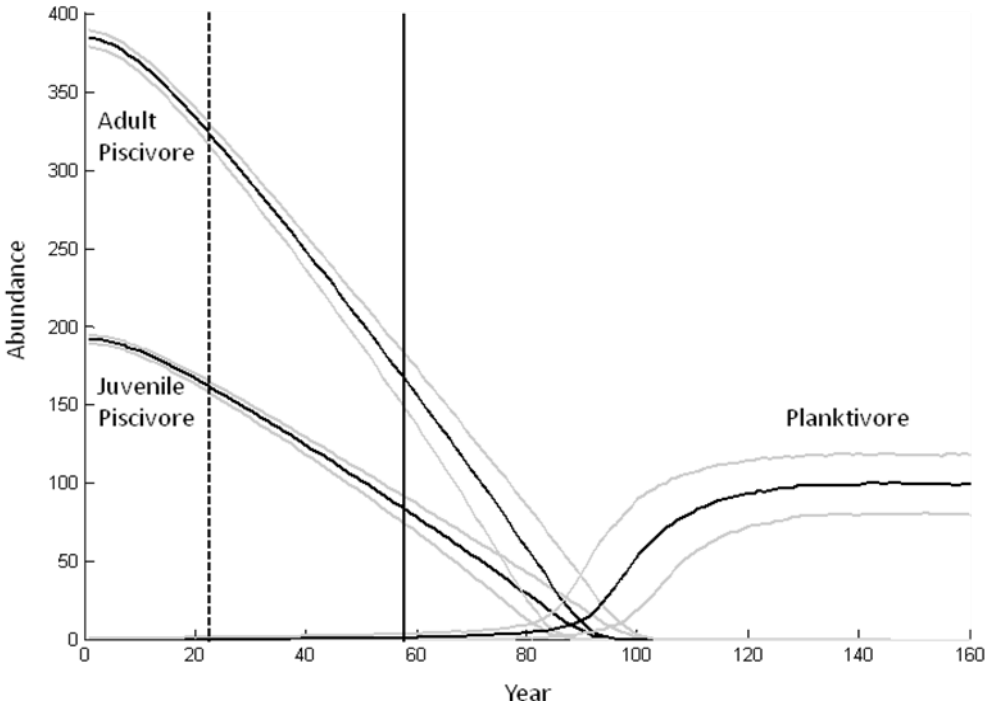

Figure 6 Time series median percentiles (black lines) and 10th and 90th percentiles (gray lines) for each group during the high noise simulations with β=1.6 (red noise). The vertical dashed line is the attractor switch point (year 23); the vertical solid line is the "point of no return" (year 58).

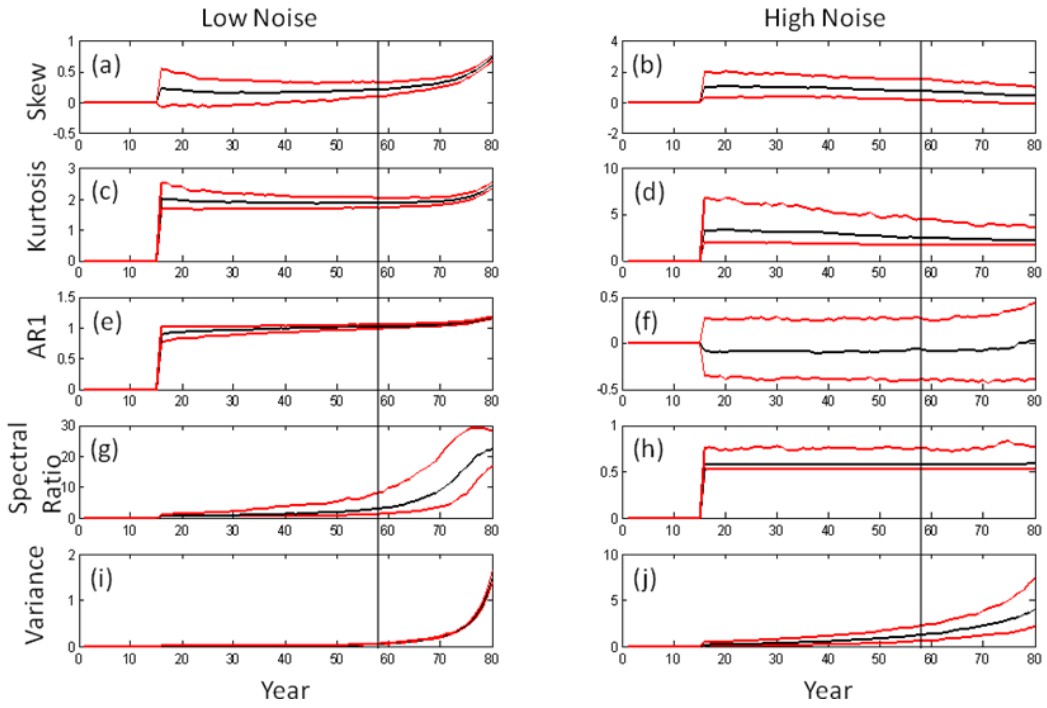

Figure 7 Indicator values vs. time calculated using annual data (with β=1.6) under low noise (left panels) and high noise (right panels) shown up to year 80. Median (black line), and 90th and 10th percentiles (red lines) of all runs shown. Indicators are calculated using a moving window of 15 years. The vertical line is the "point of no return". Skew and kurtosis indicators (a-d) show no indication of an impending shift until after the "point of no return". AR1 indicator (e,f) provides a gradual indication at low noise but fails at high noise. Spectral ratio (g,h) performs well at low noise but also fails at high noise. Variance indicator (i,j) performs well at low noise and high noise.

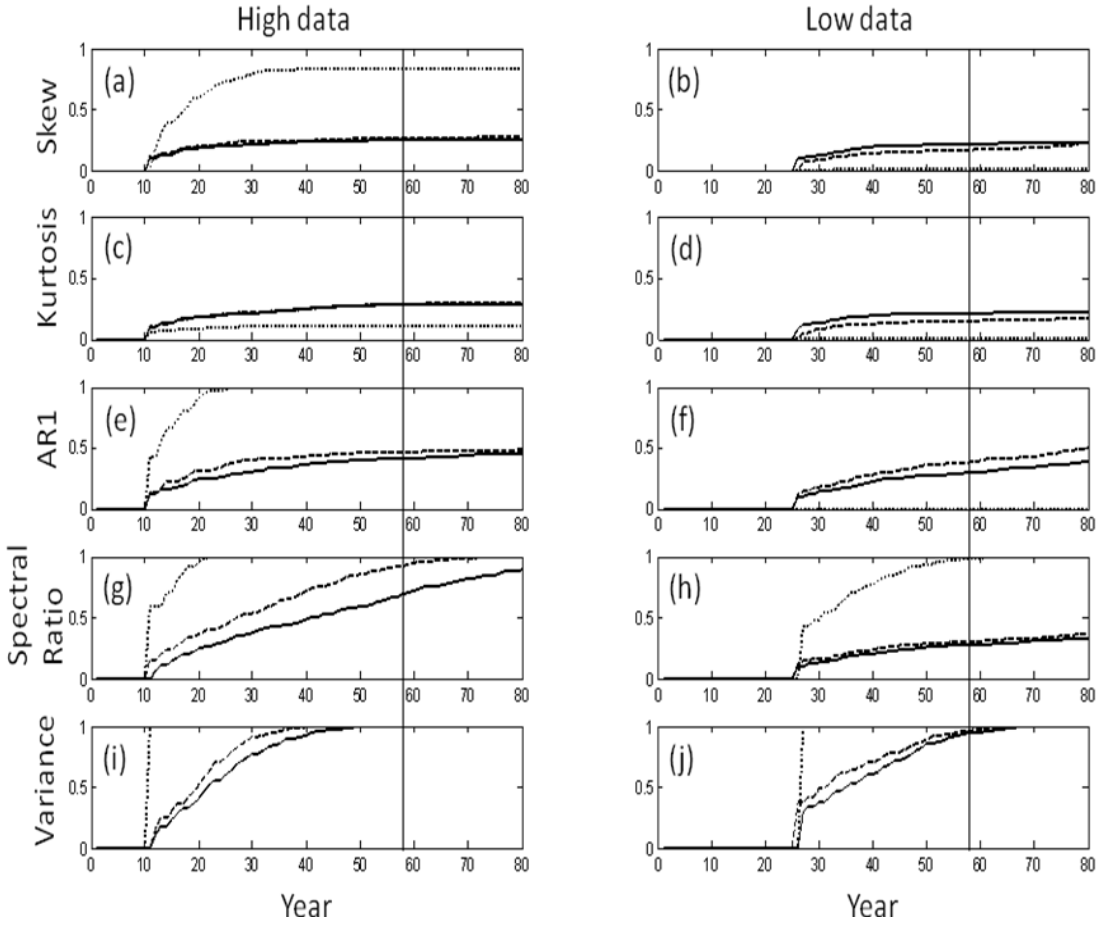

Figure 8 Probability of concluding a phase shift is occurring for all runs using high data (50 within year points, left panels) vs. low data (1 annual point, right panels) shown up to year 80. Low noise (dotted line), medium noise (dashed line) and high noise (solid line). Vertical line is the "point of no return" (year 58). High data indicators use a moving window of 10 years to calculate the slope of the indicator; low data indicators use a 15 year moving window to calculate the indicator and a 10 year moving window to calculate the slope. Skew and kurtosis (a-d) fail to detect the phase shift >70% of the time, except in the skew low noise runs. AR1 indicator performs well in the high data, low noise simulations but performs poorly elsewhere. Spectral ratio performs poorly in the low data, medium and high noise runs. Variance performs well across all treatments.

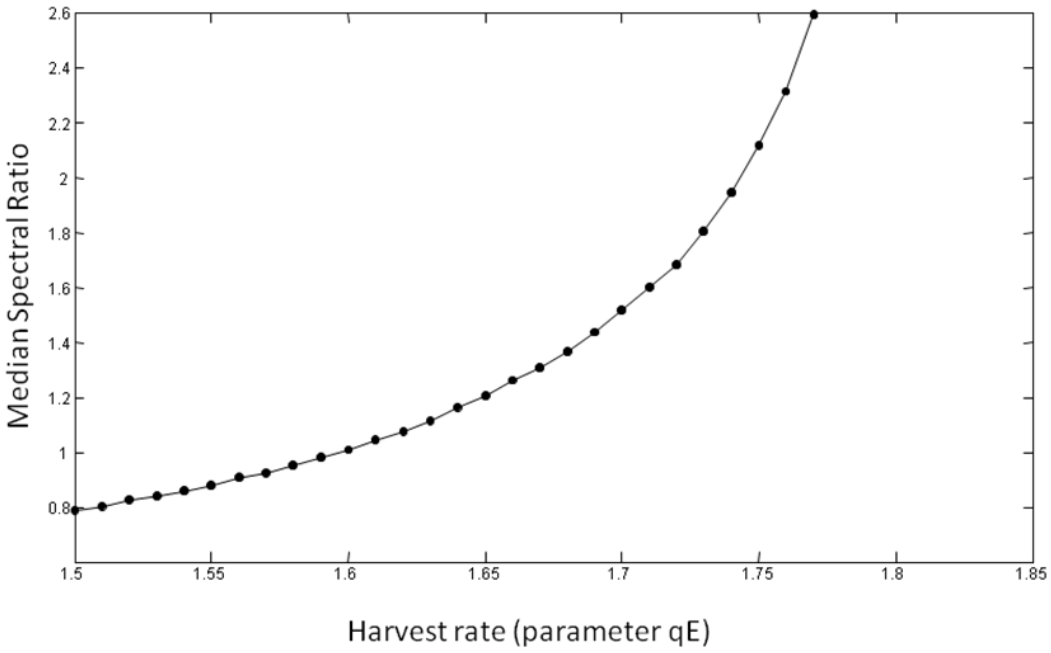

Figure 9 Median spectral ratio vs. harvest rate of adult piscivores for equilibrium simulations, i.e. no possibility of a regime shift. Results are for low noise simulations across all noise colors. The spectral ratio exceeds one for a range of harvest parameters when the system is at equilibrium, therefore a spectral ratio exceeding one is not a reliable indicator of an impending shift.